

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 10:15:55 ; Search time 49.36 Seconds
(without alignments)
7254.087 Million cell updates/sec

Title: US-09-603-124B-1
Perfect score: 1581
Sequence: 1 gcaggtaacgtccacaggt.....gacgttaagtgacaagcgag 1581

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapart 60.0
Searched: 331203 seqs, 113238999 residues

Word size : 0
Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTGUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	1.1	2289	2 US-08-463-081B-30	Sequence 30, Appl
C 2	18	1.1	2289	2 US-08-463-379A-30	Sequence 30, Appl
C 3	18	1.1	2289	2 US-08-463-300B-30	Sequence 30, Appl
C 4	18	1.1	2289	3 US-08-463-074B-30	Sequence 30, Appl
C 5	18	1.1	2289	3 US-08-463-585C-30	Sequence 30, Appl
C 6	18	1.1	2289	3 US-08-652-446-30	Sequence 30, Appl
C 7	18	1.1	2289	1 US-08-473-122-1	Sequence 1, Appl
C 8	18	1.1	2946	2 US-08-472-478-1	Sequence 1, Appl
C 9	18	1.1	2946	2 US-08-463-081B-7	Sequence 7, Appl
C 10	18	1.1	2946	2 US-08-463-379A-7	Sequence 7, Appl
C 11	18	1.1	2946	2 US-08-463-390B-7	Sequence 7, Appl
C 12	18	1.1	2946	3 US-08-463-074B-7	Sequence 7, Appl
C 13	18	1.1	2946	3 US-08-463-585C-7	Sequence 7, Appl
C 14	18	1.1	2946	3 US-08-652-446-7	Sequence 7, Appl
C 15	17	1.1	300	2 US-08-308-494A-16	Sequence 16, Appl
C 16	17	1.1	309	1 US-08-467-393-3	Sequence 3, Appl
C 17	17	1.1	322	3 US-08-434-000A-11	Sequence 11, Appl
C 18	17	1.1	322	4 US-09-312-157-11	Sequence 11, Appl
C 19	17	1.1	324	3 US-08-483-749A-7	Sequence 7, Appl
C 20	17	1.1	327	2 US-08-002-324-3	Sequence 3, Appl
C 21	17	1.1	327	5 PCT-US94-00261-3	Sequence 3, Appl
C 22	17	1.1	331	3 US-08-836-561-32	Sequence 32, Appl
C 23	17	1.1	337	4 US-09-171-945-8	Sequence 8, Appl
C 24	17	1.1	360	1 US-08-447-422-13	Sequence 13, Appl
C 25	17	1.1	384	2 US-08-656-586-1	Sequence 1, Appl
C 26	17	1.1	384	4 US-08-619-491-1	Sequence 1, Appl
C 27	17	1.1	384	5 PCT-US95-07302-1	Sequence 1, Appl

C 28	17	1.1	393	2 US-08-116-778E-39	Sequence 39, Appl
C 29	17	1.1	393	2 US-08-438-562-39	Sequence 39, Appl
C 30	17	1.1	393	2 US-08-483-528B-2	Sequence 2, Appl
C 31	17	1.1	393	2 US-08-673-798C-2	Sequence 2, Appl
C 32	17	1.1	462	1 US-07-846-421-23	Sequence 23, Appl
C 33	17	1.1	462	2 US-08-737-129A-3	Sequence 3, Appl
C 34	17	1.1	468	6 5455030-4	Patent No. 5455030
C 35	17	1.1	669	2 US-08-190-199A-66	Sequence 66, Appl
C 36	17	1.1	672	2 US-08-190-199A-62	Sequence 62, Appl
C 37	17	1.1	705	4 US-09-171-945-16	Sequence 16, Appl
C 38	17	1.1	708	2 US-08-190-199A-60	Sequence 60, Appl
C 39	17	1.1	711	2 US-08-190-199A-64	Sequence 64, Appl
C 40	17	1.1	719	3 US-08-279-772A-7	Sequence 7, Appl
C 41	17	1.1	720	1 US-08-061-092A-2	Sequence 2, Appl
C 42	17	1.1	720	3 US-08-902-486-10	Sequence 10, Appl
C 43	17	1.1	729	6 5455030-14	Patent No. 5455030
C 44	17	1.1	729	6 5455030-16	Patent No. 5455030
C 45	17	1.1	732	2 US-08-553-497A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-463-081B-30/C
Sequence 30, Application US/08463081B
Patent No. 5871560-30
Patent No. 5871560-30
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding C85 Polypeptide,
VECTOR AND TRANSFORMED CELL THEREOF, AND EXPRESSION THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: PRETTY, SCHROEDER & POPLAWSKI
STREET: 444 South Flower St. - Suite 1900
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,081B
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO. 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-463-081B-30

Query Match 1.1% Score 18; DB 2; Length 2289;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1333 tgcctgtgcttgagat 1350
DB 2113 TGCCGTGCTTGAGAT 2096

RESULT 2
US-08-461-379A-30/c
; Sequence 30, Application US/08461379A
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,390B
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108
; FILING DATE: 27-OCT-1994
; APPLICATION NUMBER: USSN 08/104,736
; FILING DATE: 10-AUG-1993
; APPLICATION NUMBER: USSN 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Anzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: DART-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)407-0700
; TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-461-379A-30

Query Match 1.1%; Score 18; DB 2; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1333 tgcctgtgcttgagat 1350
DB 2113 TGCCGTGCTTGAGAT 2096

RESULT 3
US-08-462-390B-30/c
; Sequence 30, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
; APPLICANT: Smith, K. A., & Beadling, C.
; TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
; TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
; NUMBER OF SEQUENCES: 35

Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1333 tgcctgtgcttgagat 1350
DB 2113 TGCCGTGCTTGAGAT 2096

RESULT 4
US-08-463-074B-30/c
; Sequence 30, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector an
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTI, SCHROEDER & POPLAWSKI
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,074B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:

Query Match 1.1%; Score 18; DB 2; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1333 tgcctgtgcttgagat 1350
DB 2113 TGCCGTGCTTGAGAT 2096

RESULT 4
US-08-463-074B-30/c
; Sequence 30, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector an
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTI, SCHROEDER & POPLAWSKI
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,074B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:

SUMMARIES

ALIGNMENTS

/organism="Corynebacterium glutamicum"

Query Match 100.00; Score 1581; DB 6; Length 349980;
 Best Local Similarity 100.00; Pred. No. 0;
 Matches 1581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 7066
 2,100,001 to 2,449,980"
 BASE COUNT 80724 a 98367 c 90490 g 80399 t
 ORIGIN

Qy 1 gcaggtacgcctccacggtgatgcagacatgattgcaactatcaatagccaaacac 60
 Db 182723 GCAGGTAAAGCCTCCACGCTGATTGACAGATGATTGCTCACTATCAATAGCAACAC 182664
 Qy 61 aactaaagacagcagctcaacgcaaggaatagtttaagagtgaccctccacactgga 120
 Db 182663 AACTAAAGCAGCCTCAACGCAAGGAATAGTTTAAAGGTGACCTCCACACTTGGGA 182604
 Qy 121 tcttcacagatattgattctccgcctccacactatcggtattggaggccggaat 180
 Db 182603 TTCTCACAGATATTGATCTGTGCTCCGCTCCACTCATCGGTATTGGCGAGCCGGAAT 182344
 Qy 181 gctgggtgctccgaactcctgcttcccggtggaagacagtcactggttccgattgcaa 240
 Db 182343 GTCTGGGCTTGCCTCGAATCTGCTTCCCTCCGCGGTAGACAGTCACTGTTCCGATGCCAA 182484
 Qy 241 agattccgcacactgcttccactccggtggtggaagccacatcgagtggtggaacgc 300
 Db 182483 AGATTCCGCACTTGTCTTCCACTCCGCGCGGTGGAGCCACCATCGCATGGGACAGC 182424
 Qy 301 tgcggaaaccttgagcttccgggaacttccacacgtcggtggtgacactcttttgcgc 360
 Db 182423 TCGGGAACCTTTGAGCTTTCCGCGCAACTTCCACCGCTGCTGCTGACCTTTTTCGCGC 182364
 Qy 361 cattccgcaagacacccggaaacttctgtgcacgtggaaggaagccattccggttattcg 420
 Db 182363 CATTCGCAAGAACACCGGAACCTGCTGTCGACGTGAAGAGGCAATTCGCGTTATTCG 182304
 Qy 421 tgcctccgactgtgtggcgaacttctgtagaggtccacacagctcttattcgaggtac 480
 Db 182303 TGGCTCCGACTGTTGGCGAATGCTGGAGGCTCCACCAAGCTCTTATTCGCGGTAC 182244
 Qy 481 ccaggttaagactccacacactctattctgtgtgtagctagcagcagcggtgcatgga 540
 Db 182243 CCAGGTAAGACTTCACCACTCTATGCTGTGTGTAGCTATGTCAGGAGCGGCGCATGGA 182184
 Qy 541 tccaaagcttgctatcggtggaagcgtcacaaggtggtggaacatgacacacatggaac 600
 Db 182183 TCCAAAGCTTTGCTATGCGCGGACAGCTCAACAGGCTGCGACCAATGCGACCATGGAC 182124
 Qy 601 tggtaggtcttattcgtgaagcagatgaatctgacgcatgctgctgctgctacagcc 660
 Db 182123 TGGTAGGCTCTTTATCGCTGAGCAGATGAATCTGACGCACTGCTGCTGCTACAGCC 182064
 Qy 661 aaatgtgacgtggtgcaacatgtggaacacgacacacacacacacacacacacacac 720
 Db 182063 AAATGTTGACGTGTGACCAATGTGAACACGACCACTGGACTCTTTAAACCCCTGA 182004
 Qy 721 agcctacttccaaagtggtgacagatttcgacagacacacacacacacacacacacac 780
 Db 182003 AGCCTACTTCCAAAGTGTGACCAATTTGCGAGGACGATCAACCCGAGGCGGACGCTGT 181944
 Qy 781 tgtgtgacgactctccacgacgagctgagggagggagggagggagggagggagggat 840
 Db 181943 TGTGTGCTCAACGATCTTCACGACCGGAGCTGGGGAGAGGGTCTGCTCGCGAGGGTAT 181884
 Qy 841 caagactgttggttaccgctgacgcccgtacagggcaccacacacacacacacacacac 900
 Db 181883 CAAGACTGTTGGTTAGGTTAGCCTGACCGCTGACAGGCGACACCTGTGAGGTTCCAGCGAT 181824
 Qy 901 ggtaccactgttgatttcccaagttgtcgaggaagggcaccggtgacacacacacacac 960

Db 181823 GGCTACCATCTGATTCCTCAAGTTGTGCGACAGGACCGCGCCCACTCAACATCGA 181764
 Qy 961 tggacaggaagatctctgattctctcaaatccctgggtgacacagtgactcaacgggtgc 1020
 Db 181763 TGGACAGGAAGTATCTGTGATTCTTCAATCTCTGGTGATCATATGTTACTCAACGGTGC 181704
 Qy 1021 agcgcctctgctggcgggatacctcgtggtggtggtggtggtggtggtggtggtggt 1080
 Db 181703 AGCGCCCTGCTGGCGGATACCTGCTGGGTGGGACGTCGACAGCTTGTGTGAAGCTTT 181644
 Qy 1081 gtcgagattcttcggggtgctgacgcgcctcttgatttcacaggtgctatcgaggggcgga 1140
 Db 181643 GTCGGAATTTCTCGGGGTGCGACCCCTTTGTTTCCACGGTGTATCGAGGGCGGCA 181584
 Qy 1141 atttaattggcgtgctatttattgattgattacacacacacacacacacacacacac 1200
 Db 181583 ATTTAATGGCGCTGCTATTATTATGATGATTAACACACACACACACACACACACACAC 181524
 Qy 1201 gctgcagcgtgctgcgcacccgggtggaagccgctggaaagggcgtgctatcgtcgctt 1260
 Db 181523 GCTCAGCGCTGCGCGCACCGCGGTGAAGGCCGCTGGAAAGGCCGCTGTCATCGTGCCTT 181464
 Qy 1261 ccaacacatttatactcacgcacacacacacacacacacacacacacacacacacac 1320
 Db 181463 CCAACACATTTATCTCACACACACATAGATTTCCAAAGGAGTTTCGCGGGGCGACTGTC 181404
 Qy 1321 actgagagagcgtgcgcgtggtggtggtggtggtggtggtggtggtggtggtggtggt 1380
 Db 181403 ACTGGCAGCGCTGCGGTGCTGTGATGATTTACGGAGCGCGGAAACCGGTGGATGG 181344
 Qy 1381 cgtgctcctcggaactcaccacacacacacacacacacacacacacacacacacac 1440
 Db 181343 CGTGTCTCGGAATCTACACGATGCGATGACCATTCACAGTGTGTACGACCTAATTT 181284
 Qy 1441 cctcagctccagacagcattgacgaacacacacacacacacacacacacacacacac 1500
 Db 181283 CTCTGATCTCCAGAACGCTATGCGAAATCGAGGACCTTAATGACATCGCTGCTCACCAT 181224
 Qy 1501 gggctcaggttccgtgac 1560
 Db 181223 GGGTCAGAGTTCCGTGACCATGCTTCTCCAGAACTCTGGATCTGGATCTGCAAAACATTA 181164
 Qy 1561 ggcgttaagtgaaacagggcag 1581
 Db 181163 GGACGTAAGTGAACAGGCGAG 181143

RESULT 2

AB015023
 LOCUS AB015023 2291 bp DNA BCT 06-FEB-1999
 DEFINITION Corynebacterium glutamicum genes for MurC and FtsQ, complete cds.
 ACCESSION AB015023
 VERSION AB015023.1 GI:3688934
 KEYWORDS FtsQ; MurC
 SOURCE Corynebacterium glutamicum DNA.
 ORGANISM Corynebacterium glutamicum
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 (bases 1 to 2291)
 Wachi,M.
 Direct Submission
 Submitted: (27-MAY-1998) to the DDBJ/EMBL/GenBank databases, Masaki Wachi, Tokyo Institute of Technology, Department of Bioengineering, 4259 Nagatsuta Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail:mwachi@bio.titech.ac.jp, Tel:81-45-924-5770, Fax:81-45-924-5820)
 REFERENCE 2 (sites)
 Wachi,M., Wajayathana,C.D., Teraoka,H. and Nagai,K.
 A murC gene from coryneform bacteria
 Appl. Microbiol. Biotechnol. (1998) in press
 Location/Qualifiers

102b

Qy 1568 agtgaacaggcag 1581
 |||||
 Db 1561 ACTGAACAGGCAG 1574

RESULT 3

BLA242646 4622 bp DNA BCT 19-JAN-2000
 Brevibacterium lactofermentum murD (partial), ftsW, murG and murC
 genes, located upstream of ftsQ-ftsZ.

ACCESSION AJ242646
 VERSION AJ242646.1 GI:6723455
 KEYWORDS ftsW gene; ftsW protein; murC gene; murC protein; murD gene; murD
 protein; murG gene; murG protein.

SOURCE Corynebacterium glutamicum.
 ORGANISM Corynebacterium glutamicum.
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;

REFERENCE 1 (bases 1 to 4622)
 Ramos, A., Honrubia, P. and Gil, J. A.
 Characterization and chromosomal organization of the murD-murC
 region of Brevibacterium lactofermentum ATCC 13669
 unpublished

JOURNAL 2 (bases 1 to 4622)
 Gil, J. A.
 Direct Submission
 Submitted (17-MAY-1999) Gil J. A., Microbiology, University of Leon,
 Campus de Vegazana s/n, 24071 Leon, SPAIN
 Location/Qualifiers

1. 4622
 /organism="Corynebacterium glutamicum"
 /strain="ATCC 13669"
 /db_xref="taxon:1718"

1. 333
 /gene="murD"
 /feature="peptidoglycan biosynthesis"

1. 333
 /codon_start=1
 /transl_table=1
 /evidence=experimental

/product="MurD protein"
 /protein_id="CA866322.1"
 /db_xref="GI:6723456"

/translation="LKADIAPLVKKHEQRIKALVLGADRAEIVAAKLEHAFQASVF
 VTKTEFPFMEIEIVTENSISEPDVTLAPAAASLDMFKMGQKGDLPZAHNIIGTI
 KGLUTEKG"

336..1808
 /gene="ftsW"
 336..1808
 /gene="ftsW"

/function="cell division"
 /codon_start=1
 /transl_table=1
 /evidence=experimental

/product="ftsW protein"
 /protein_id="CA866323.1"
 /db_xref="GI:6723457"

/translation="MTWGLSKKEDAPVTCATKTRTGLIRERISGNWDLARPUTDI
 MILCTVYLCLQVWYVSSMTSLREGSVNGVAVRQGIHVLGPFAMVALMTRP
 OTTERLSNLYLVILLAVQICIGTGREZVGSQSNIALGPIPOPSFAKVAIAV
 WGLTLAGKGPQVWNLNLPFGVGPAPNAPLITMGDGHANGSPVYLWFLMFPAG
 TAWKIALGVLLIALLAVLNGGPFSSRFYITDALPGRHFDVRIAPVQYGLS
 LAGSGVGLGQSKRMFTLPEARNDFIPAIIGBELGLAGLWIALFAGLLYGLR
 TAKSHDPTGLMAATLITASVVSQAFNIQYVGLPLVPTGIQLPMISAGOTSAILTLA
 SMILLISCARHEPTEVAMASYGRPAIDRLGLREPSSTLTTSNABLSNKNKAKQK
 SESSEIVSGPLWRACDRTPSACATKSNASNTIGSSATLW"

1994..3112
 /gene="murG"
 1994..3112
 /gene="murG"

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

gene
 CDS

/function="peptidoglycan biosynthesis"
 /codon_start=1
 /transl_table=1
 /product="MurG protein"
 /protein_id="CA866324.1"
 /db_xref="GI:6723458"

/translation="MKSSELRIMANSKPMRVVAGGCTAGCTEPAALVAELRDKH
 GATVAGTARGLETTLPORGFEHLLEFPVVPNPNMDLLKLPFRVACALQDARKA
 LKDDAAGVCGGCTVSPATMAKSLGIPFYZHARAGANKLVKLGQVGNVAV
 KQSGHDDVYGIPFIVLISGARDESAAARDTWGLDKQKQRTIPVGGSGSVSNKA
 TGAIDQVZGQVGHVGRKHELPAKPGTHPPVPIIDDMQAATVADLIIVCRSGRM
 TVAEYTAAGVPALIVPLPHNGSQALNAQAVIKAGARIDDDADDTAOTLIDATLIDIL
 LNPSTHSDMSDAKATSTAGNASTVIADNIAATINSQHN"

3147..4607
 /gene="murC"
 3147..4607
 /gene="murC"

/function="peptidoglycan biosynthesis"
 /codon_start=1
 /transl_table=1
 /product="MurC protein"

/protein_id="CA866325.1"
 /db_xref="GI:6723459"

/translation="MTPHLOSADQIDLSRVHLIGIGAGAGSVARILLARGKTTGGS
 DAKDSRTLPRAVGATIANVGAENLELSGELPTVTVTSFNAIPDNPELYRAREZO
 IPVIRSDLLGLLESTOVLIAGTHGKTSTSSVVAHQAGDQBSFAIGQLKAG
 TNAHKGVEFTAEADESDASLIRKQVAVVNEPDLQFTPEAKTQVDDPAG
 RITPNGLVCLNDPDAELGERSVRKCIKTVGYCTFRVQNHFEVPAKATVDSQVY
 AGGTATVINTDQGVYVILQIDQHWLQKQALAGLITNGDQVHFKVYVYDQVY
 RPEFENADQCKPAGLITVQDYNHPTVAVVCRICRKAAGRGVRIYAPQPHLT
 ERTMEQKESASLADAVLEIYGVSEQPVYVGSSEIITDANTIPVYEPNFSAV
 PERIAEIPNDIVLTWAGSVTHLAPLIDLOQNN"

989 a 1256 c 1317 g 1062 t
 BASE COUNT
 ORIGIN

Query Match 97.88; Score 1545.6; DB 1; Length 4622;
 Best Local Similarity 98.88; Pred. No. 0;
 Matches 1557; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Oy 1 gcaggtaacgctccacgggtgattgcagacatgatgtcgaactatcaatagccacac 60
 |||||
 Db 3047 GCAGGTAACTCCACGGTATTCAGACATGATTGCTGCAACTATCAATAGCCACAC 3106

Oy 61 asctaaacgacacagctcaacgcaaggaagatttaagggtgacacccacacttggg 120
 |||||
 Db 3107 AACTAAACGACACGCTCAACGCAAGGAATAGTTTAAAGGTGACCACTCCACACTTGA 3166

Oy 121 tctgcacagatatattgactctcccgccctccacccatcgcgtatttggcggagccgg 180
 |||||
 Db 3167 TTCGCAAGATATTTGATCTGCGCGCTCCACCTCATCGGTATTCGCGAGCCGAA 3226

Oy 181 gctcgggttcggcgaatccctgttgcgcggtgaagacagtcactcgttccgagtgcaa 240
 |||||
 Db 3227 GTCTGCGCTTGGCCGAAATCTGCTTGCCTCCGCGGTGAAGACAGTCACTGTTCGATGCCAA 3286

Oy 241 agattccgcacatttcttcactctcgcgcggtggagccacatcgagtggggacgc 300
 |||||
 Db 3287 AGATTCGCCACCTTGTCTCCACTCGCGCGGTGGGAGCCACCATCGCATGGGACAGC 3346

Oy 301 tgcggaaacacttgactctccggcggaacttcccaocgctcgttggacaccttcttgcgc 360
 |||||
 Db 3347 TCGGGAACACTTGAGCTTTCGCGGAACTTCCACCGCTCGCTGCTCTTTTGGCGC 3406

Oy 361 ctttccgcaagaacacccggaacttcttcgctcgcgctgaagagcattccgggtattcg 420
 |||||
 Db 3407 CATTCGCCAAGACACCCGGAACCTTGTTCGTCACCTGACAGAGCAATTCGCGTTATTCG 3466

Oy 421 tgcctccgactctgttggcggaattctgtgaagacctccaccagcttcttgcgggtac 480
 |||||
 Db 3467 TCGCTCCGATCTGTTCGGCGAATTCTCGGAAGGCTCCACCCAGGTCTTGTATCCCGGTAC 3526

Oy 481 ccacggtgaagacccctccaccctctctctgtctgttgcgtatgcagggcagcggtgca 540

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: February 15, 2002, 09:58:11 ; Search time 1598.42 Seconds
(without alignments)
16317.389 Million cell updates/sec

Title: US-09-603-124B-1

Perfect score: 1581

Sequence: 1 gcaggtaacgctcccgct.....gacgtatgtgacacagcgag 1581

Scoring table: OLIGO-MTC

Gapop 60.0 , Gapext 50.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.om.*

20: em.ov.*

21: em.pat.*

22: em.ph.*

23: em.pl.*

24: em.ro.*

25: em.sts.*

26: em.sy.*

27: em.un.*

28: em.vi.*

29: em.yi.*

30: em.htgo.hum.*

31: em.htgo.in.*

32: em.htgo.ro.*

33: em.htgo.hum.*

34: em.htgo.in.*

35: em.htgo.ro.*

36: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1581	100.0	349980	6	AX127150	Sequence
2	1458	92.2	1458	6	AX122446	Sequence
3	1122	71.0	2291	1	AB015023	Sequence
4	1050	68.4	1458	6	E28467	Novel murc
5	951	60.2	4622	1	BLA242646	Novel murc
6	236	14.9	4116	1	AB003132	Sequence
7	236	14.9	5546	1	BLF752	Sequence
8	63	4.0	1116	6	AX122447	Sequence
9	21	2.1	1073	8	AF022730	Oryza sat
10	21	1.3	6843	2	AC019805	Drosophila
11	21	1.3	85875	2	AC010010	Drosophila
12	21	1.3	126039	9	AC012665	Homo sapi
13	21	1.3	137569	2	AC084093	Homo sapi
14	21	1.3	163027	2	AC022846	Homo sapi
15	21	1.3	174822	9	AC012375	Homo sapi
16	21	1.3	175590	2	AF307159	Homo sapi
17	21	1.3	181848	2	AC090199	Homo sapi
18	21	1.3	310120	3	AB003468	Drosophila
19	20	1.3	234	8	TA8111570	Sequence
20	20	1.3	1253	6	AX122173	Sequence
21	20	1.3	1253	6	AX122688	Sequence
22	20	1.3	1253	6	AX122688	Sequence
23	20	1.3	1253	6	AX122688	Sequence
24	20	1.3	1253	6	AX122688	Sequence
25	20	1.3	1253	6	AX122688	Sequence
26	20	1.3	1253	6	AX122688	Sequence
27	19	1.2	11994	8	AF169016	Halobacter
28	19	1.2	15000	8	AF169016	Halobacter
29	19	1.2	32274	1	SCC8A	Streptococcus
30	19	1.2	56917	1	AME16952	Streptococcus
31	19	1.2	67480	2	AF279874	Streptococcus
32	19	1.2	85217	8	ATT17133	Streptococcus
33	19	1.2	104770	9	HSJ245M18	Streptococcus
34	19	1.2	109612	9	AL512503	Human DNA
35	19	1.2	114856	9	HSJ245M18	Human DNA
36	19	1.2	116501	2	AC080803	Homo sapi
37	19	1.2	119060	2	AC023390	Homo sapi
38	19	1.2	119456	2	AC084789	Homo sapi
39	19	1.2	131047	2	AC002421	Homo sapi
40	19	1.2	139232	2	AL353762	Homo sapi
41	19	1.2	144727	9	HSJ245M18	Human DNA
42	19	1.2	145382	2	AC073223	Homo sapi
43	19	1.2	148260	2	AL596222	Homo sapi
44	19	1.2	148882	2	AC018381	Homo sapi
45	19	1.2	150436	2	AC044119	Homo sapi

ALIGNMENTS

RESULT 1	AX127150	349980 bp	DNA	PAT	11-MAY-2001
LOCUS	Sequence 7066 from Patent EP1108790.				
DEFINITION	AX127150 AX114121				
ACCESSION	AX127150.1 GI:114041138				
VERSION					
KEYWORDS	Corynebacterium glutamicum.				
SOURCE	Corynebacterium glutamicum.				
ORGANISM	Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae; Actinomycetales; Corynebacteriales; Corynebacteriaceae; Corynebacterium;				
REFERENCE	1 (bases 1 to 349980)				
AUTHORS	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochikawa, K., Yokoi, H., Tetsu, H., Senoh, A., Ikeda, M. and Ozaki, A.				
TITLE	Novel polynucleotides				
JOURNAL	Patent EP 1108790-A 7066 20-JUN-2001;				
FEATURES	LOCUS: KYOKA HAKKO KOGYO CO., LTD. (JP)				
SOURCE	Location/Qualifiers				
	1. 349980				

[illegible]

Op 501, 5^{ds} 43

```

/ gene="fts3"
/ codon_start=1
/ transl_table=11
/ product="cell division protein"
/ protein_id="CAA70158.1"
/ db_xref="GI:1769661"
/ db_xref="SWISS-PROT:P94337"
/ translation="MTSPNNTLAKIKVYGVGGGVNAVRNIEGLKGVETIVNTDTS
QALMFSDADVKLIGRATKRGKGLGAGANPMGRASADHKNEIETIKGADVPVYVTA
GGTGTGMAPVAGIAGKRGATLIGVYTFPEPGRRTROABESIAALKKVCVDTLII
IPDRLLEGDANLIMREFAEDVHNGVQITNLITPGVIMDFADKMEVNAAL
SKLVESGSGARDNRVVSATDNLNGLSPILSRATMDGATGVLLSPAGSDGLMEVNAAL
SVKRESDVNLKIFCTIITDNLGDEVTVIATGDAASAPASRRBAGIPAPAPAS
PVQGVPTNATLPPEKESIEGGARENDPVLSEAGASHIETTSOGGLFTTGNDR
EDREREDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDR
2553 3295
/ gene="y41n"
/ gene="y41n"
/ gene="y41n"
/ gene="y41n"
/ gene="y41n"
/ note="hypothetical protein similar to YFif protein"
/ codon_start=1
/ transl_table=11
/ protein_id="CAA70159.1"
/ db_xref="GI:1769662"
/ db_xref="SPTRMBL:P94338"
/ translation="MDSLOPNRVRVFRVTRAGVSSQSPYASPNLDHVGDDPQAVA
SHRNLIATIGLSDFGVYVIMQIHSHVITIDEAPAGQAVEATDLVTIQRGLAVAL
LVADVPVGLTSDAGVTAAVHAGRAGRNIVAKTIAMKELGAKPSRIHALGAGAA
SCANTYEPVANARDVEAKLPGLSTARTKTGTGLDIRAGLLQMLSLGVQMDISDPCT
TEDDTLFTFRREGTTQRQAGVVMPLKEA"
3376 .4041
/ note="ORF3"
/ codon_start=1
/ transl_table=11
/ product="putative YPj3 protein"
/ protein_id="CAS866326.1"
/ db_xref="GI:6723446"
/ translation="MGAIRIDLEHNRPESGVLLPTFKFHPVEDIKTLQELGVTA
GRENQERAKALEPQDPMFKIGOSKANSIARWAAVHSVDSEKIALAGRGVA
LATEREDVAGLEPCFIOLSDPDSRGRTPLSCVTLQALCKISDTHLRFELGKCVPP
LGHDPKAFSQARDVLGSLGEHPDRSLERSNGHSDLIATIKHGSITVRVGTGLNLR
PLA"
4118 .4576
/ note="ORF6"
/ codon_start=1
/ transl_table=11
/ product="putative YAK 1 protein"
/ protein_id="CAS866327.1"
/ db_xref="GI:6723447"
/ translation="MSMKTKTEFFGLAPYCAEHADYVADPRYSGTAYAPAYETPR
DTGAPAPAPAPSPAPSPQSTVYVFLSHSFSDAQYTGGAQPDGDAVYVDMSLLSR
SEARRIVFAGSLGAFALRGKQKIDSVFVAVVPELSNISTSEERANIR"
BASE COUNT 1317 a 1479 c 1520 g 1230 t
ORIGIN

Query Match 14.9% score 236; DB 1; Length 5546;
Best Local Similarity 100.0%; Mismatches 0; Ident. 0; Gaps 0;
Matches 236; Conservative 0;

QY 1346 gagattacggagcgcaacaacccggcgagtagcgtagctctcgcgaatcatcacccgat 1405
DB 49 GAGATTTACGAGCGCCGCAACACCGGTGATCGCGCTGTCCTCGGAATCATCACCCGAT 108
QY 1406 gcgatgacattcccgtagctagcaactaatcttcgcaggtccgcgaacgcatgca 1465
DB 109 GCGATGACCATTCGAGTGGTGTACGAACTAATTTCTTCGAGTGCCTCCGAGACCGCATGCA 168
QY 1466 gaatcgaggaactaingtacatcgctgcctcaactagggtagcaggttccgtgacatgctt 1525
DB 169 GAATTCGAGGACCTAATACATCGTGTCTCACCATGGGTGCGAGGTCCGTCCGACACATGCTT 228

```

Db 229 GCTCCAGAAATCTGTCAGCTGCAAAACAATAGGACGTAAGTGAACAGGCAG 284
|||||

RESULT 8
AX122447 1116 bp DNA PAT 11-MAY-2001
LOCUS
DEFINITION Sequence 3363 from Patent EP1108790.
ACCESSION AX122447
VERSION AX122447.1 GI:14039694
KEYWORDS
SOURCE Corynebacterium glutamicum.
ORGANISM Corynebacterium glutamicum

REFERENCE
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 2363 20-JUN-2001;
KTOMA HAKKO KOGYO CO., LTD. (JP)
FEATURES
Location/Qualifiers
1. 1116
/organism="Corynebacterium glutamicum"
/db_xref="taxon:1718"

BASE COUNT 234 a 317 c 319 g 246 t
ORIGIN
Query Match 4.04; Score 63; DB 6; Length 1116;
Best Local Similarity 100.04; Pred. No. 4.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gcaggtacgctccaggtgattgcagacattgattgcacattcatcaatgacacac 60
|||||
Db 1054 GCAGGTACGCTCCAGGTGATTGCAGACATTGATTGCACATTATCAATGACCAAC 1113
Qy 61 aac 63
Db 1114 AAC 1116

RESULT 9
AF022730 1073 bp mRNA PLN 30-OCT-1997
LOCUS
DEFINITION Oryza sativa glyceraldehyde-3-phosphate dehydrogenase subunit
(GAPDH) mRNA, partial cds.
ACCESSION AF022730
VERSION AF022730.1 GI:2570494
KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzedeae; Oryza.
TITLE Isolation and characterization of glyceraldehyde-3-phosphate
dehydrogenase from rice
JOURNAL Unpublished
2 (bases 1 to 1073)
AUTHORS Lee, M.C., Kim, C.S. and Eun, M.Y.
DIRECT SUBMISSION
TITLE Submitted (04-SEP-1997) Cyto Genetics, National Institute of
Agricultural Science and Technology, RDA, Sedun-Dong, Suwon
441-707, Korea
FEATURES
Location/Qualifiers
1. 1073
/organism="Oryza sativa"
/strain="Milyang 23"
/db_xref="taxon:4530"
/tissue_type="seed"

Query Match 1.34; Score 21; DB 2; Length 6843;
Best Local Similarity 100.04; Pred. No. 7.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/dev_stage="immature (milky stage)";
/clone="KCDM2189";
1. >1073
/gene="GAPDH"
3
/size ">1073"
/accession="GAPDH"
/codon_start=1
/product="glyceraldehyde-3-phosphate dehydrogenase subunit"
/protein_id="AA82131.1"
/db_xref="GI:2570494"
/translation="MSPCPSPPPCHSRGGGLDFGLRSSSLPLRDATSDNPNP
VFRTHAVTNGGARRAPTEAKLVKVAINGFRGRNPLSCWHGRGDDSPPLAIPINDT
GGVQASRLKYNSTGLFDADVPKVPENPLSVEYIKVSDRNPSPNLMWGLGIEL
LIEGTGLFFNAGAKHFOAGAKVLITAPGKGDIPITYVGVNADQVSPDEPIGNAS
CTTTCCLAPFKLDKFGIINGTMTTHTSYTGDRVFDGSHRDLARARAAALNPTTS
TGAALVALVPLKGLKGLIALRVPTNVSVDLVAVQVSKRLAEVYVQAFRDSAAH
ELKASSRER"
BASE COUNT 196 a 380 c 335 g 162 t
ORIGIN

Query Match 1.34; Score 21; DB 8; Length 1073;
Best Local Similarity 100.04; Pred. No. 9.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1076 ggcttgctgaggttctccggc 1096
|||||
Db 93 GGCTTGCTGAGATTCTCCGCC 113

RESULT 10
AC019805/C 6843 bp DNA HTG 03-JAN-2000
LOCUS
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC019805
VERSION AC019805.1 GI:6665092
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 6843)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDW:10211017 by the submitter.
For more information on this record e-mail to flyce@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 6843
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 1966 a 1466 c 1449 g 1962 t
ORIGIN

Query Match 1.34; Score 21; DB 2; Length 6843;
Best Local Similarity 100.04; Pred. No. 7.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 262 actccggccgtggagccac 282
|||||
Db 5976 ACTCCGGCCGTGGAGCCAC 5956

RESULT 11
AC010010

AC010010 89875 bp DNA 18-FEB-2000
 Drosophila melanogaster clone RPC198-1376, *** SEQUENCING IN
 PROGRESS ***, 53 unordered pieces.
 AC010010
 AC010010.3 GI:6996694
 HTG: HTGS_PHASE1.
 SOURCE
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 89875)
 Musny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
 Bodota,B., Bouck,J.J., Bowie,S., Brooks,A., Buhal,C., Bunac,C.,
 Burkett,C., Burrows,J., Carter,M., Chacto,J., Chen,Z., Cox,C.,
 David,R., Delgado,O., Deshaio,D., Ding,X., Fomah-Rashid,N.,
 Duran-Rocha,S., Durbin,K.J., Fernandez,C., Farragudo,D.,
 Forcum-Taney,J., Frantz,P., Ganesh,R., Correll,J.H., Gorrell,L.L.,
 Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hughes,M.,
 Kellay,S., Kondejewski,N., Kong,I., Kovar,C., Leal,B., Li,S.,
 Holloway,C., Hosak,K., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
 Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,J.,
 Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
 Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
 Osvald,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
 Cui,M., Reiter,D., Rivas,M., Samuel,S., Say,J., Scherer,S.,
 Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Suqiang,R.,
 Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wahbah,M.,
 Washington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
 Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
 Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 89875)
 Worley,K.C.
 Direct Submission
 Submitted (11-SEP-1999) Ruman Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Feb 18, 2000 this sequence version replaced gi:5981484.
 ***** Genome Center *****
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ***** Project Information *****
 Center project name: DRBG
 Center clone name: RPC198-1376
 ***** Summary Statistics *****
 Sequencing vector: M13; 108821
 Chemistry: Dye-Primer; BigDye; 684 of reads
 Assembly program: Phrap; version 0.980611
 Consensus quality: 41466 bases at least Q40
 Consensus quality: 57245 bases at least Q30
 Consensus quality: 64739 bases at least Q20
 Estimated insert size: 75108; sum-of-contigs estimation
 Quality coverage: 0.8x in Q20 bases; sum-of-contigs estimation
 ***** NOTE: This is a 'working draft' sequence. It currently
 * consists of 53 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1256: contig of 1256 bp in length
 * 1257 1276: gap of unknown length
 * 1277 2355: contig of 1079 bp in length
 * 2356 2375: gap of unknown length
 * 2376 3453: contig of 1077 bp in length
 * 3453 3472: gap of unknown length

3472 4275: contig of 803 bp in length
 4276 4293: gap of unknown length
 4296 3562: contig of 1267 bp in length
 3563 3582: gap of unknown length
 3583 3777: contig of 1195 bp in length
 3778 3821: gap of unknown length
 3822 3861: contig of 1264 bp in length
 3862 3903: gap of unknown length
 3904 3973: gap of 972 bp in length
 3974 3994: contig of 861 bp in length
 3995 3995: gap of unknown length
 3996 10766: contig of 812 bp in length
 10767 10786: gap of unknown length
 10787 11616: contig of 830 bp in length
 11617 11636: gap of unknown length
 11637 12464: contig of 828 bp in length
 12465 12484: gap of unknown length
 12485 13645: contig of 1161 bp in length
 13646 13665: gap of unknown length
 13666 14089: contig of 824 bp in length
 14090 14090: gap of unknown length
 14091 15861: contig of 1352 bp in length
 15862 15921: contig of 1040 bp in length
 15922 16941: gap of unknown length
 16942 18465: contig of 1524 bp in length
 18466 18485: gap of unknown length
 18486 19316: contig of 831 bp in length
 19317 19336: gap of unknown length
 19337 20903: contig of 1567 bp in length
 20904 20923: gap of unknown length
 20924 22150: contig of 1227 bp in length
 22151 23295: contig of 1125 bp in length
 23296 23315: gap of unknown length
 23316 24330: contig of 995 bp in length
 24331 26237: gap of unknown length
 26238 27815: contig of 1378 bp in length
 27816 27835: gap of unknown length
 27836 28794: contig of 959 bp in length
 28795 28814: gap of unknown length
 28815 30095: contig of 1381 bp in length
 30096 30115: gap of unknown length
 30116 31085: contig of 970 bp in length
 31086 31105: gap of unknown length
 31106 32390: contig of 1285 bp in length
 32391 34245: contig of 1835 bp in length
 34246 35631: contig of 1366 bp in length
 35632 35831: gap of unknown length
 35832 37039: contig of 1408 bp in length
 37040 38607: gap of unknown length
 38608 38627: gap of unknown length
 38628 40420: contig of 1793 bp in length
 40421 41701: contig of 1261 bp in length
 41702 43313: contig of 1598 bp in length
 43314 43339: gap of unknown length
 43340 45717: contig of 2378 bp in length
 45718 45737: gap of unknown length
 45738 47273: contig of 1536 bp in length
 47274 47293: gap of unknown length
 47294 48893: contig of 1599 bp in length
 48894 48913: gap of unknown length
 48914 50344: contig of 1432 bp in length
 50345 52143: gap of unknown length
 52144 50365: contig of 1779 bp in length

AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. MO 63108, USA
COMMENT On Jul 3, 2001 this sequence version replaced gi:13794751.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 11391
 Center clone name: 2277_K12
 ----- Summary Statistics
 Sequencing Vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly Program: Phrap; version 0.960731
 Consensus quality: 13629 bases at least Q40
 Consensus quality: 137130 bases at least Q30
 Consensus quality: 137294 bases at least Q20
 Insert size: 139000; agarose-fp
 Quality coverage: 137369; sum-of-contigs
 Quality coverage: 7.8 in Q20 bases; agarose-fp
 Quality coverage: 7.9 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 37849: contig of 37849 bp in length
 * 37850 37949: gap of 100 bp
 * 37950 82422: contig of 44473 bp in length
 * 82423 83522: gap of 100 bp
 * 83523 137569: contig of 55047 bp in length.
 * Location/Qualifiers
 * 1. 137569
 * /organism="Homo sapiens"
 * /db_xref="taxon:9606"
 * /chromosome="8"
 * /map="8"
 * /clone_lib="CIVD Human BAC"
 * /note="assembly_fragment"
 * misc_feature
 * 37950 82422
 * /note="assembly_fragment"
 * misc_feature
 * 82523 137569
 * /note="assembly_fragment"
 * BASE COUNT 42095 a 26887 c 26625 g 41757 t 201 others
 * ORIGIN
 *
 * Query Match 1.3%; Score 21; DB 2; Length 137569;
 * Best Local Similarity 100.0%; Pred. No. 6;
 * Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 *
 * QY 863 gctgacgcctcagcagcac 883
 * |||||
 * Db 33227 GCTGACGCGTACAGGCAC 33207
 *
 * RESULT 14
 * AC022846
 * LOCUS
 * DEFINITION Homo sapiens chromosome 8 clone RP11-253N21 map 8, WORKING DRAFT
 * SEQUENCE, 6 unordered pieces.
 * AC022846
 * ACCESSION
 * AC022846.3 GI:13121397
 * VERSION
 * KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 * SOURCE
 * human
 * ORGANISM
 * Homo sapiens
 * Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 * Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 * 1 (bases 1 to 163027)
 * REFERENCE
 * Birren,B., Linton,L., Musbaum,C., Lander,E., Abraham,H., Allen,N.,
 * Anderson,S., Baldwin,J., Barne,N., Becklerly,R., Beda,F.,
 * Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Castle,A.,
 * Choapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 * Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 * Ferrera,P., FitzHugh,W., Forrest,C., Gage,D., Galegan,J.,
 * Gardyna,S., Grant,G., Hagos,B., Hesford,A., Horton,L.,
 * Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 * Landers,T., Lehotsky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
 * MacDonald,P., Marquis,N., McEwan,P., McGuire,A., McKernan,K.,
 * McPheters,R., Meldrum,J., Menusil,L., Morrow,J., Naylor,J.,
 * Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Paterson,K.,
 * Piarre,N., Pisani,C., Pollara,Y., Raymond,C., Riley,R., Rothman,D.,
 * Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 * Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 * Tizabi,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 * Zimmer,A. and Zody,N.
 * Direct Submission
 * Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
 * Research, 320 Charles Street, Cambridge, MA 02141, USA
 * On Feb 25, 2001 this sequence version replaced gi:1139770.
 * All repeats were identified using RepeatMasker:
 * Smit, A.F.A. & Green, P. (1996-1997)
 * http://ftp.genome.washington.edu/RM/RepeatMasker.html
 * ----- Genome Center
 * Center: Whitehead Institute/ MIT Center for Genome Research
 * Center code: WBIR
 * Web site: http://www.seq.wi.mit.edu
 * Contact: sequence_submissions@genome.wi.mit.edu
 * ----- Project Information
 * Center project name: 16161
 * Center clone name: 253.N.21
 * ----- Summary Statistics
 * Sequencing Vector: N13; M7815; 39% of reads
 * Chemistry: Dye-terminator Big Dye; 100% of reads
 * Assembly Program: Phrap; version 0.960731
 * Consensus quality: 160669 bases at least Q40
 * Consensus quality: 161671 bases at least Q30
 * Consensus quality: 162136 bases at least Q20
 * Insert size: 157000; agarose-fp
 * Insert size: 162327; sum-of-contigs
 * Quality coverage: 13.6 in Q20 bases; agarose-fp
 * Quality coverage: 13.1 in Q20.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 23967 24067: contig of 23967 bp in length
 * 24068 24087: gap of 100 bp
 * 24088 25128: contig of 1061 bp in length
 * 25129 25228: gap of 100 bp
 * 25229 24037: contig of 8809 bp in length
 * 34038 34137: gap of 100 bp
 * 34138 36813: contig of 22676 bp in length
 * 56814 56913: gap of 100 bp
 * 56914 92224: contig of 35311 bp in length
 * 92225 92324: gap of 100 bp
 * 92325 163027: contig of 70703 bp in length.
 * Location/Qualifiers
 * 1. 163027
 * /organism="Homo sapiens"
 * /db_xref="taxon:9606"
 * /chromosome="8"
 * /map="8"
 * /clone_lib="RP11-253N21"
 * /clone_lib="RP11-253N21"
 * /note="assembly_fragment"
 * misc_feature
 * 1. 23967
 * /note="assembly_fragment"

AUTHORS

Birren,B., Linton,L., Musbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barne,N., Becklerly,R., Beda,F.,
 Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Castle,A.,
 Choapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
 Ferrera,P., FitzHugh,W., Forrest,C., Gage,D., Galegan,J.,
 Gardyna,S., Grant,G., Hagos,B., Hesford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Landers,T., Lehotsky,J., Levine,R., Liu,C., Liu,G., Locke,K.,
 MacDonald,P., Marquis,N., McEwan,P., McGuire,A., McKernan,K.,
 McPheters,R., Meldrum,J., Menusil,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Paterson,K.,
 Piarre,N., Pisani,C., Pollara,Y., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Tizabi,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,N.
 Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 25, 2001 this sequence version replaced gi:1139770.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WBIR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 16161
 Center clone name: 253.N.21
 ----- Summary Statistics
 Sequencing Vector: N13; M7815; 39% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly Program: Phrap; version 0.960731
 Consensus quality: 160669 bases at least Q40
 Consensus quality: 161671 bases at least Q30
 Consensus quality: 162136 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 162327; sum-of-contigs
 Quality coverage: 13.6 in Q20 bases; agarose-fp
 Quality coverage: 13.1 in Q20.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 6 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 23967 24067: contig of 23967 bp in length
 24068 24087: gap of 100 bp
 24088 25128: contig of 1061 bp in length
 25129 25228: gap of 100 bp
 25229 24037: contig of 8809 bp in length
 34038 34137: gap of 100 bp
 34138 36813: contig of 22676 bp in length
 56814 56913: gap of 100 bp
 56914 92224: contig of 35311 bp in length
 92225 92324: gap of 100 bp
 92325 163027: contig of 70703 bp in length.
 Location/Qualifiers
 1. 163027
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone_lib="RP11-253N21"
 /clone_lib="RP11-253N21"
 /note="assembly_fragment"
 misc_feature
 1. 23967
 /note="assembly_fragment"

```

clone_end:SP6
vector_side:left
24068..25128
/note="assembly_fragment"
25259..34037
/note="assembly_fragment"
34138..56813
/note="assembly_fragment"
56914..92224
/note="assembly_fragment"
92325..163027
/note="assembly_fragment"
clone_end:T7
vector_side:right
BASE COUNT 48042 a 32608 c 32206 g 49671 t 500 others
ORIGIN
Query Match 1.3% Score 21; DB 2; Length 163027;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 863 gctgacgccctacagccac 883
DB 90055 GCTGACGCCCTACAGCCAC 90075

RESULT 15
AC013275
LOCUS AC013275 174822 bp DNA PRI 09-AUG-2001
DEFINITION Homo sapiens BAC clone RP11-464C8 from 2, complete sequence.
ACCESSION AC013275
VERSION AC013275.8 GI:15144471
KEYWORDS BAC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 174822)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome research. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE
2 (bases 1 to 174822)
AUTHORS Sun, H., Hakenson, W. and Connors, J.
TITLE The sequence of Homo sapiens BAC clone RP11-464C8
JOURNAL Unpublished
REFERENCE
3 (bases 1 to 174822)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
4 (bases 1 to 174822)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 9, 2001 this sequence version replaced gi:13270787.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.edu
-----
Summary Statistics
-----
Center project name: R_MH0464C08
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

```

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.T. and de Jong, P.J. (1998) 'An improved
approach for construction of bacterial artificial chromosome
libraries, Genomics 31:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-393J17; the clone sequenced
to the right is AC016076. Actual start of this clone is at base
position 1 of RP11-464C8; actual end is at base position 174822 of
RP11-464C8.

The region between 76913 to 76922 is covered only by a single ml3
subclone. Digest information confirms the assembly.

FEATURES

```

source
1..174822
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-464C8"
1305..1548
/clone_lib="RP11-11"
/repeat_family="MER2_type"
1864..1934
/repeat_family="MIR"
1935..2234
/repeat_family="Alu"
2255..2334
/repeat_family="MIR"
3496..4429
/misc_feature
/note="similar to EST AL589988 (NID:913368825)"
5380..5666
/repeat_family="Alu"
6188..6501
/repeat_family="Alu"
6975..6998
/repeat_family="(TTTG)n"
8647..8809
/repeat_family="Alu"
8823..8884
/repeat_family="MER2_type"
9261..9396
/repeat_family="MER1_type"
9357..9711
/repeat_family="Alu"
9712..9738
/repeat_family="MER1_type"
10174..10321
/repeat_family="Alu"
10322..10366
/repeat_family="(CA)n"
10367..10453
/repeat_family="Alu"
10467..10527

```

```

repeat_region /rpt_family="MERL_type"
repeat_region 10604..10930
repeat_region /rpt_family="ERV1"
repeat_region 10931..10942
repeat_region /rpt_family="Alu"
repeat_region 10943..10974
repeat_region /rpt_family="(TG)n"
repeat_region 10975..11271
repeat_region /rpt_family="Alu"
repeat_region 11272..11429
repeat_region /rpt_family="ERV1"
repeat_region 11430..11925
repeat_region /rpt_family="ERV1"
repeat_region 12011..12116
repeat_region /rpt_family="Alu"
repeat_region 12117..12763
repeat_region /rpt_family="ERV1"
repeat_region 12764..12798
repeat_region /rpt_family="MALR"
repeat_region 12799..13097
repeat_region /rpt_family="Alu"
repeat_region 13098..13493
repeat_region /rpt_family="MALR"
repeat_region 13494..13832
repeat_region /rpt_family="MALR"
repeat_region 13833..14250
repeat_region /rpt_family="ERV1"
repeat_region 14251..15302
repeat_region /rpt_family="MALR"
repeat_region 15303..15912
repeat_region /rpt_family="MALR"
repeat_region 15913..16092
repeat_region /rpt_family="ERV1"
repeat_region 16093..17059
repeat_region /rpt_family="MERL_type"
repeat_region 17060..18107
repeat_region /rpt_family="MERL_type"
repeat_region 18108..18385
repeat_region /rpt_family="ERV1"
repeat_region 18386..19126
repeat_region /rpt_family="Alu"
misc_feature 19127..20210
repeat_region /note="similar to EST AA668276 (NID:92629775) ab78ell.s1"
repeat_region 20211..20938
repeat_region /rpt_family="MIR"
repeat_region 20939..21202
repeat_region /rpt_family="L2"
repeat_region 21203..21643
repeat_region /rpt_family="L2"
repeat_region 21644..21968
repeat_region /rpt_family="L2"
repeat_region 21969..22662
repeat_region /rpt_family="ERV1"
repeat_region 22663..22871
repeat_region /rpt_family="ERV1"
repeat_region 22872..23362
repeat_region /rpt_family="MERL_type"
repeat_region 23363..23833
repeat_region /rpt_family="MIR"
repeat_region 23834..24465
repeat_region /rpt_family="Alu"
misc_feature 24466..24802
misc_feature /note="similar to EST A1631785 (NID:94683115) wa36a10.x1"
misc_feature 24803..24885
misc_feature /note="similar to EST BF589995 (NID:g11682319)"
misc_feature 24886..25228
misc_feature /note="CpG island (GC=73.5, o/e=0.80, CpG=89)"
misc_feature 25229..25228
misc_feature /note="similar to EST BF532525 (NID:g11619888)"
misc_feature 25229..25228
misc_feature /note="similar to EST BF787117 (NID:g12092153)"
misc_feature 25229..25228
misc_feature /note="similar to EST A1507063 (NID:g4404914) vj32h01.x1"

```

```

misc_feature 25180..25228
misc_feature /note="similar to EST AN742372 (NID:g7654159) up56c03.y1"
repeat_region 26043..26175
repeat_region /rpt_family="MIR"
repeat_region 26345..26419
repeat_region /rpt_family="L2"
repeat_region 27624..27751
repeat_region /rpt_family="MIR"

```

```

Query Match 1.38; Score 21; DB 9; Length 174822;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1537 cctggatcagctgcaaaacaa 1557
DB 6112 cctggatcagctgcaaaacaa 6133

```

Search completed: February 15, 2002, 10:47:55
Job time: 3984 sec

```

/gene'fts5'
/function='division initiation protein or cell division
protein'
/note='alternative gene name div8*'
/codon_start=1
/transl_table=11
/protein_id='CAA70161.1'
/db_xref='GI:1769960'
/db_xref='SPTRXBL:P94336'
/translation='MKKVIYLYGVVVVYVAILGVVAVFPVLKVGNIETGATRTD
PQIVLVSIGVSEGENLFVDNATAGQNIYELPMKSVTVNRAIPSTIVETLTEREPA
FIKRAQDGRVITDEGRILLTGTPGVTVSYSGADGENSEYLPVAVIATVNAIKRQDAQM
TESIQVCEAPDQFDILLKNDRIYNGSSSENNHDKAVANSTVLKREQGQRNINISSPM
VTVR'
1211..2527
/gene'fts5'
1211..2527

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 10:15:55 ; Search time 49.36 seconds
(without alignments)
7256.087 Million cell updates/sec

Title: US-09-603-124B-1

Perfect score: 1581
Sequence: 1 gcaggttaacgctccacggt.....gacgttaagtaacagggcag 1581 -

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702405

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2.6/ptodata/2/lna/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/lna/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/lna/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/lna/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/lna/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	1.1	2289	2 US-08-463-081B-30	Sequence 30, Appl
C 2	18	1.1	2288	2 US-08-461-379A-30	Sequence 30, Appl
C 3	18	1.1	2289	2 US-08-462-390B-30	Sequence 30, Appl
C 4	18	1.1	2289	3 US-08-463-074B-30	Sequence 30, Appl
C 5	18	1.1	2289	3 US-08-465-585C-30	Sequence 30, Appl
C 6	18	1.1	2289	3 US-08-652-446-30	Sequence 30, Appl
C 7	18	1.1	2946	1 US-08-473-122-1	Sequence 1, Appl
C 8	18	1.1	2946	2 US-08-472-478-1	Sequence 1, Appl
C 9	18	1.1	2946	2 US-08-463-081B-7	Sequence 7, Appl
C 10	18	1.1	2946	2 US-08-461-379A-7	Sequence 7, Appl
C 11	18	1.1	2946	2 US-08-462-390B-7	Sequence 7, Appl
C 12	18	1.1	2946	3 US-08-463-074B-7	Sequence 7, Appl
C 13	18	1.1	2946	3 US-08-465-585C-7	Sequence 7, Appl
C 14	18	1.1	2946	3 US-08-652-446-7	Sequence 7, Appl
C 15	17	1.1	3009	1 US-08-308-494A-16	Sequence 16, Appl
C 16	17	1.1	322	3 US-08-467-393-3	Sequence 3, Appl
C 17	17	1.1	322	3 US-08-434-000A-11	Sequence 11, Appl
C 18	17	1.1	322	4 US-09-312-157-11	Sequence 11, Appl
C 19	17	1.1	324	3 US-08-463-149A-7	Sequence 7, Appl
C 20	17	1.1	327	2 US-08-002-124-3	Sequence 3, Appl
C 21	17	1.1	327	3 PCT-US93-00261-3	Sequence 3, Appl
C 22	17	1.1	357	3 US-08-836-361-32	Sequence 32, Appl
C 23	17	1.1	357	4 US-09-171-945-8	Sequence 8, Appl
C 24	17	1.1	360	1 US-08-447-422-13	Sequence 13, Appl
C 25	17	1.1	384	2 US-08-656-586-1	Sequence 1, Appl
C 26	17	1.1	384	4 US-08-619-491-1	Sequence 1, Appl
C 27	17	1.1	384	5 PCT-US93-07302-1	Sequence 1, Appl

Sequence 39, Appl
Sequence 39, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 25, Appl
Sequence 3, Appl
Patent No. 5453030
Sequence 66, Appl
Sequence 62, Appl
Sequence 16, Appl
Sequence 60, Appl
Sequence 64, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 10, Appl
Patent No. 5453030
Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-463-081B-30/c
; Sequence 30, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960-5837487
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
; NUMBER OF INVENTION: Vector and Transformed Cell Thereof, and Expression Ther.
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; STREET: 444 South Flower St. - Suite 1900
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,081B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,736
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviane Ansel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: 965 39150 (DART-060)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-463-081B-30

Query Match 1.1%; Score 18; DB 2; Length 2289;

RESULT 3
US-08-62-3908-30/c
Sequence 30, Application US/0862390B
Patent No. 5887894
GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadling, C.
TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
TITLE OF INVENTION: Transformed Call Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 35

APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector and
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
ADDRESSEE: (B) STREET: 444 South Flower St. - Suite
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993
PRIOR APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Anzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-463-074B-30

Query Match 1.18; Score 18; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1333 tgcctgtctctgagat 1350
|||||
DB 2113 TCCCTGTCTCTGAGAT 2096

RESULT 5
US-08-465-585C-30/c
Sequence 30, Application US/08465585C
Patent No. 6057314
GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadling, C.
TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vector and Expression Thereof
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: US 07/796,066
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 20-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Anzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 base pairs
TYPE: nucleic acid

Query Match 1.18; Score 18; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1333 tgcctgtctctgagat 1350
|||||
DB 2113 TCCCTGTCTCTGAGAT 2096

RESULT 5
US-08-465-585C-30/c
Sequence 30, Application US/08465585C
Patent No. 6057314
GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadling, C.
TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vector and Expression Thereof
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: US 07/796,066
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 20-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Anzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-463-585C-30

Query Match 1.18; Score 18; DB 3; Length 2289;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1333 tgcctgtctctgagat 1350
|||||
DB 2113 TCCCTGTCTCTGAGAT 2096

RESULT 6
US-08-652-446-30/c
Sequence 30, Application US/08652446
Patent No. 6057427
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide, Vector and Transformed Cell Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 38
CURRENT APPLICATION DATA:
FILING DATE: 08/08/652,446
APPLICATION NUMBER: US/08/652,446
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP App. # 96921319.8
FILING DATE: 5-JAN-1998
APPLICATION NUMBER: PCT/US/96/09194
FILING DATE: 5-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,108
FILING DATE: 2-OCT-1994
APPLICATION NUMBER: 08/463,074
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,337
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,390
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,585
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,081
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,379
FILING DATE: 5-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/739,523
FILING DATE: 29-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Anzel, Ph. D.
REGISTRATION NUMBER: 30,930

REFERENCE/DOCKET NUMBER: P566 40035
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 622-7700
 TELEFAX: (213) 489-4210
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2289 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-652-446-30

Query Match 1.18; Score 18; DB 3; Length 2289;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2113 TGCCTGGTCTTGAGAT 2096
 DB 2113 TGCCTGGTCTTGAGAT 2096

RESULT 7
 US-08-473-122-1/c
 Sequence 1, Application US/08473122
 Patent No. 5652340
 GENERAL INFORMATION:
 APPLICANT: KOHMI-SHIGEMATSU, TERUMI
 APPLICANT: KOHMI, YOSHINORI
 APPLICANT: DICKINSON, LILIANE A
 TITLE OF INVENTION: MATRIX-ASSOCIATING DNA-BINDING PROTEIN.
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME AND METHODS FOR DETECTING
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CAMPBELL AND FLORES
 STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
 CITY: SAN DIEGO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,122
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/934,034
 FILING DATE: 21-AUG-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: CAMPBELL, CATHRYN
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 9293
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619 535-9001
 TELEFAX: 619 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-473-122-1

Query Match 1.18; Score 18; DB 1; Length 2946;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 TGCCTGGTCTTGAGAT 1350
 DB 2327 TGCCTGGTCTTGAGAT 2310

RESULT 9
 US-08-463-081b-7/c
 Sequence 7, Application US/08463081b
 Patent No. 5871960
 GENERAL INFORMATION:
 APPLICANT: Smith, Kendall A. & Beadling, Carol
 TITLE OF INVENTION: Nucleic Acids Encoding CR3 Polypeptide,
 TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression There

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 TGCCTGGTCTTGAGAT 1350
 DB 2327 TGCCTGGTCTTGAGAT 2310

RESULT 8
 US-08-472-478-1/c
 Sequence 1, Application US/08472478
 Patent No. 5869621
 GENERAL INFORMATION:
 APPLICANT: KOHMI-SHIGEMATSU, TERUMI
 APPLICANT: KOHMI, YOSHINORI
 APPLICANT: DICKINSON, LILIANE A
 TITLE OF INVENTION: MATRIX-ASSOCIATING DNA-BINDING PROTEIN.
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME AND METHODS FOR DETECTING
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CAMPBELL AND FLORES
 STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
 CITY: SAN DIEGO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,478
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/934,034
 FILING DATE: 21-AUG-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: CAMPBELL, CATHRYN
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 9293
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619 535-9001
 TELEFAX: 619 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-472-478-1

Query Match 1.18; Score 18; DB 2; Length 2946;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 TGCCTGGTCTTGAGAT 1350
 DB 2327 TGCCTGGTCTTGAGAT 2310

RESULT 9
 US-08-463-081b-7/c
 Sequence 7, Application US/08463081b
 Patent No. 5871960
 GENERAL INFORMATION:
 APPLICANT: Smith, Kendall A. & Beadling, Carol
 TITLE OF INVENTION: Nucleic Acids Encoding CR3 Polypeptide,
 TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression There

1 NUMBER OF SEQUENCES: 35
 1 CORRESPONDENCE ADDRESS:
 1 ADDRESSEE: A. PETTI, SCHROEDER & POPLAWSKI
 1 STREET: 44 South Flower St. - Suite 1900
 1 CITY: Los Angeles
 1 STATE: California
 1 COUNTRY: USA
 1 ZIP: 90071
 1 COMPUTER READABLE FORM:
 1 MEDIUM TYPE: Floppy disk
 1 COMPUTER: IBM PC compatible
 1 OPERATING SYSTEM: PC-DOS/MS-DOS
 1 SOFTWARE: Patent in Release #1.0,
 1 CURRENT APPLICATION DATA:
 1 APPLICATION NUMBER: US/08/463,081B
 1 FILING DATE: 5-JUN-1995
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: US 08/104,736
 1 FILING DATE: 10-AUG-1993
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: US 07/796,066
 1 FILING DATE: 20-NOV-91
 1 ATTORNEY/AGENT INFORMATION:
 1 NAME: Viviana Anzel, Ph. D.
 1 REGISTRATION NUMBER: 30,930
 1 REFERENCE/DOCKET NUMBER: P66 36150 (DART-060)
 1 TELEPHONE: (213) 622-7700
 1 TELEFAX: (213) 489-4210
 1 INFORMATION FOR SEQ ID NO: 7:
 1 SEQUENCE CHARACTERISTICS:
 1 TYPE: nucleic acid
 1 LENGTH: 2946 base pairs
 1 STRANDEDNESS: single
 1 TOPOLOGY: linear
 1 MOLECULE TYPE: cdna
 1 FEATURE:
 1 NAME/KEY: CDS
 1 LOCATION: 215..2503
 1 US-08-463-081B-7
 1
 1 Query Match 1.1%, Score 18; DB 2; Length 2946;
 1 Best Local Similarity 100.0%, Pred. No. 12;
 1 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1
 1 QY 1333 tgcctgtgtcttgagat 1350
 1 |||||
 1 DB 2327 TGCCTGTGCTTGAGAT 2310
 1
 1 RESULT 10
 1 US-08-461-379A-7/c
 1 Sequence 7, Application US/08461379A
 1 Patent No. 5871961
 1 GENERAL INFORMATION:
 1 APPLICANT: Smith, Kendall A. & Beadling, Carol
 1 TITLE OF INVENTION: Nucleic Acids Encoding CRS Polypeptide,
 1 TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
 1 TITLE OF INVENTION: Expression Thereof
 1 NUMBER OF SEQUENCES: 35
 1 CORRESPONDENCE ADDRESS:
 1 ADDRESSEE: Ratner & Prestia
 1 CITY: Valley Forge
 1 STATE: Pennsylvania
 1 COUNTRY: USA
 1 ZIP: 19482
 1 COMPUTER READABLE FORM:
 1 MEDIUM TYPE: Floppy disk
 1 COMPUTER: IBM PC compatible
 1 OPERATING SYSTEM: PC-DOS/MS-DOS
 1 SOFTWARE: Patent in Release #1.0,
 1 CURRENT APPLICATION DATA:
 1 APPLICATION NUMBER: US/08/463,081B
 1 FILING DATE: 5-JUN-1995
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: US 08/104,736
 1 FILING DATE: 10-AUG-1993
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: US 07/796,066
 1 FILING DATE: 20-NOV-91
 1 ATTORNEY/AGENT INFORMATION:
 1 NAME: Viviana Anzel, Ph. D.
 1 REGISTRATION NUMBER: 30,930
 1 REFERENCE/DOCKET NUMBER: P66 36150 (DART-060)
 1 TELEPHONE: (213) 622-7700
 1 TELEFAX: (213) 489-4210
 1 INFORMATION FOR SEQ ID NO: 7:
 1 SEQUENCE CHARACTERISTICS:
 1 TYPE: nucleic acid
 1 LENGTH: 2946 base pairs
 1 STRANDEDNESS: single
 1 TOPOLOGY: linear
 1 MOLECULE TYPE: cdna
 1 FEATURE:
 1 NAME/KEY: CDS
 1 LOCATION: 215..2503
 1 US-08-463-081B-7

1 SOFTWARE: Patent in Release #1.0,
 1 CURRENT APPLICATION DATA:
 1 APPLICATION NUMBER: US/08/461,379A
 1 FILING DATE: 5-JUNE-1995
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: USSN 08/330,108; 08/104,736
 1 FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
 1 ATTORNEY/AGENT INFORMATION:
 1 NAME: Viviana Anzel, Ph. D.
 1 REGISTRATION NUMBER: 30,930
 1 REFERENCE/DOCKET NUMBER: DART-070
 1 TELECOMMUNICATION INFORMATION:
 1 TELEPHONE: (610)470-0700
 1 TELEFAX: (610)470-0701
 1 INFORMATION FOR SEQ ID NO: 7:
 1 SEQUENCE CHARACTERISTICS:
 1 LENGTH: 2946 base pairs
 1 TYPE: nucleic acid
 1 STRANDEDNESS: single
 1 TOPOLOGY: linear
 1 MOLECULE TYPE: cdna
 1 FEATURE:
 1 NAME/KEY: CDS
 1 LOCATION: 215..2503
 1 US-08-461-379A-7
 1
 1 Query Match 1.1%, Score 18; DB 2; Length 2946;
 1 Best Local Similarity 100.0%, Pred. No. 12;
 1 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1
 1 QY 1333 tgcctgtgtcttgagat 1350
 1 |||||
 1 DB 2327 TGCCTGTGCTTGAGAT 2310
 1
 1 RESULT 11
 1 US-08-462-390B-7/c
 1 Sequence 7, Application US/08462390B
 1 Patent No. 5882894
 1 GENERAL INFORMATION:
 1 APPLICANT: Smith, K. A. & Beadling, C
 1 TITLE OF INVENTION: Nucleic Acids Encoding CRS Polypeptide, Vector and
 1 TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
 1 NUMBER OF SEQUENCES: 35
 1 CORRESPONDENCE ADDRESS:
 1 ADDRESSEE: Ratner & Prestia
 1 CITY: Valley Forge
 1 STATE: Pennsylvania
 1 COUNTRY: USA
 1 ZIP: 19482
 1 COMPUTER READABLE FORM:
 1 MEDIUM TYPE: Floppy disk
 1 COMPUTER: IBM PC compatible
 1 OPERATING SYSTEM: PC-DOS/MS-DOS
 1 SOFTWARE: Patent in Release #1.0, Version #1.25
 1 CURRENT APPLICATION DATA:
 1 APPLICATION NUMBER: US/08/462,390B
 1 FILING DATE: 5-JUNE-1995
 1 PRIOR APPLICATION DATA:
 1 APPLICATION NUMBER: USSN 08/330,108
 1 FILING DATE: 27-OCT-1994
 1 APPLICATION NUMBER: USSN 08/104,736
 1 FILING DATE: 10-AUG-1993
 1 APPLICATION NUMBER: USSN 07/796,066
 1 FILING DATE: 20-NOV-91
 1 ATTORNEY/AGENT INFORMATION:
 1 NAME: Viviana Anzel, Ph. D.
 1 REGISTRATION NUMBER: 30,930
 1 REFERENCE/DOCKET NUMBER: DART-040

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610)407-0700
 TELEFAX: (610)407-0701
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 215..2503
 US-08-462-3908-7

Query Match 1.18; Score 18; DB 2; Length 2946;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 tgcctgtgcttgagat 1350
 DB 2327 tgcctgtgcttgagat 2310

RESULT 12
 US-08-463-074B-7/c
 Sequence 7, Application US/08463074B
 Patent No. 6020155
 GENERAL INFORMATION:
 APPLICANT: Smith, Kendall A. & Beadling, Carol
 TITLE OF INVENTION: Nucleic Acids Encoding CR1 Fusion Protein, Vector an
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0,
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,074B
 FILING DATE: 5-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/104,736
 FILING DATE: 10-AUG-1993
 APPLICATION NUMBER: US 07/796,066
 FILING DATE: 20-NOV-91
 ATTORNEY/AGENT INFORMATION:
 NAME: Viviana Anzel, Ph. D.
 REGISTRATION NUMBER: 30,930
 REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-4210
 TELEFAX: (213) 489-4210
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 215..2503
 US-08-463-074B-7

Query Match 1.18; Score 18; DB 3; Length 2946;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 tgcctgtgcttgagat 1350
 DB 2327 tgcctgtgcttgagat 2310

RESULT 13
 US-08-465-585C-7/c
 Sequence 7, Application US/08465585C
 Patent No. 6027914
 GENERAL INFORMATION:
 APPLICANT: Smith, K. A., & Beadling, C.
 TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vecto
 TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 900071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465,585C
 FILING DATE: 5-JUNE-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USN 08/330,108
 FILING DATE: 27-OCT-1994
 APPLICATION NUMBER: USN 08/104,736
 FILING DATE: 10-AUG-1993
 APPLICATION NUMBER: USN 07/796,066
 FILING DATE: 20-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Viviana Anzel, Ph. D.
 REGISTRATION NUMBER: 30,930
 REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 622-7700
 TELEFAX: (213) 4894210 7;
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 215..2503
 US-08-465-585C-7

Query Match 1.18; Score 18; DB 3; Length 2946;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1333 tgcctgtgcttgagat 1350
 DB 2327 tgcctgtgcttgagat 2310
 RESULT 14
 US-08-652-446-7/c

Sequence 7, Application US/08652446
Patent No. 6057427
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CDS
TITLE OF INVENTION: Polypeptide, Vector and Transformed Cell Thereof, and
TITLE OF INVENTION: Expression Thereof
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,446
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP APP. # 9691319.8
FILING DATE: 5-JAN-1998
APPLICATION NUMBER: PCT/US/96/09194
FILING DATE: 5-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,108
FILING DATE: 27-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,074
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,337
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,390
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,585
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,081
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,379
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/739,523
FILING DATE: 29-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Ariel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: FP66 40035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2946 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 213..2503
US-08-652-446-7

Query Match 1.1% Score 18, DB 3, Length 2946,
Best Local Similarity 100.0%, Pred. No. 12,
Matches 18, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
QY 1333 tgcggtggtgctgagct 1350
|||||
DB 2327 tccggtggtgctgagct 2310
|||||
RESULT 15
US-08-308-494A-16/c
Sequence 16, Application US/08308494A
Patent No. 5959083
GENERAL INFORMATION:
APPLICANT: Bosalet, Klaus
APPLICANT: Seeman, Gerhard
TITLE OF INVENTION: Tetraivalent Bispesific Receptors, The
TITLE OF INVENTION: Preparation and Use Thereof
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,494A
FILING DATE: 21-SEP-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,739
FILING DATE: 01-JUN-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P4118120.4
FILING DATE: 03-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kulik, David J.
REGISTRATION NUMBER: 36,576
REFERENCE/DOCKET NUMBER: 05552-1186-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-308-494A-16

Query Match 1.1% Score 17, DB 2, Length 300,
Best Local Similarity 100.0%, Pred. No. 42,
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
QY 24 tgcagacatgctctc 40
|||||
DB 30 TGCAGACATGCTCTG 14
|||||

Search completed: February 15, 2002, 10:46:25
Job time: 1831 sec

SDS

03-AUG-2001
segment 96/162.

```

Location/Qualifiers
1. .37386
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="Y270"
/complement(51. .854)
/gene="cyoQ"
/complement(51. .854)
/gene="cyoQ"
/notes="K9V131c, (MTCY270.37), len: 267, cyoQ, equivalent to M. leprae CYSO_MYCLE P46776 cyoQ protein homolog. (289 aa). Contains inositol monophosphatase family signature 1 (PS00629), significance uncertain. FASTA blast: CYSO_MYCLE P46776 cyoQ protein homolog. (289 aa) opt: 1374 z-score: 1597.0E(1) 0: (77.3% identity in 364 aa overlap)."
/codon_start=1
/tranal_table=11
/product="CyOQ"
/protein_id="CAB08639.1"
/db_xref="GI:2104347"
/db_xref="SWISS-PROT:O06244"
/translation="MYSAPAPDLTDLTDAELAAKGLLLQVRAETGFDQPT  
LEAGDQKANSLLRLRLARPGDAVLSEAHDDLRKSDRVYIDPLDTGTFSTP  
GRDWAHLHLMRRNSQKBITDAAVAPARGNVTYVTSGAAPGAPVPGTGLIA  
VSATRPAPVRLRRTLLAIOPVSIIGSGAKAMVDDGYTDAVTHAGGOWENDSAAAP

```


QY 377 cgggaactgtctgacagtggaaggaacttcgggttatttgctgctcagatctgttg 436
DB 19392 CCCGACCTGTCGAGAGCGAGCGCGCGGATTCCTCGTGTGCTGCGCGCGCGCTGTC 19333
QY 437 ggcgaactgtcgaagagctccacacgggtcttgatcgagtggtacacagcgaagaccc 496
DB 19332 GCCAATTGATGGCGGGCGGACACATGTGATGTCTACCGGCGACGACGCGACGACG 19273
QY 497 accaactctatgtctgtatgactatgcagggcagcgggcgtatggtatcaagctttgtatc 556
DB 19272 ACGAGTCATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19213
QY 557 ggcgaactgtcgaagagctccacacgggtcttgatcgagtggtacacagcgaagaccc 616
DB 19212 GCGGTGAGCTGGGGAGCGCGGATTCACGACGCGGATTCACGACGCGGATTCGTC 19153
QY 617 gctgaagcagatggaattgaagcagatcgctgctgctgctgctgctgctgctgctgct 676
DB 19152 GCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 19093
QY 677 accaactgtggaacacgac 736
DB 19092 ACCAATCATGAGTGCATGATCCTGATGATGATGATGATGATGATGATGATGATG 19033
QY 737 ttcgaagatttgcaggaagcagcagcagcagcagcagcagcagcagcagcagcagcag 796
DB 19032 TTCGACTCTCTCGGAGCGGATTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 18973
QY 797 cctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 856
DB 18972 CCCGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 18914
QY 857 ggtacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 916
DB 18913 CCGGTGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 18854
QY 917 tccaaagtgtgcaggaagcagcagcagcagcagcagcagcagcagcagcagcagcag 976
DB 18853 CCGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18794
QY 977 g-tgattcttcac 1035
DB 18793 GATGCGGTGTCGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18734
QY 1036 cggatcactgt 1095
DB 18733 CCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 18674
QY 1096 cgtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1155
DB 18673 AGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18613
QY 1156 tatttatgattgattgattgattgattgattgattgattgattgattgattgattgattg 1215
DB 18614 --TGTTCGATGCTACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18557
QY 1216 caccggtgtgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1275
DB 18556 CATGGTGTCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18497
QY 1276 ctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1335
DB 18496 TTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 18437
QY 1336 cgt 1395
DB 18436 GTTCGTACTGACGCTCTACGAGCTCTGAGAGCTCTGAGAGCTCTGAGAGCTCTGAG 18377
QY 1396 catcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1455
DB 18376 CGTCCGTGAGCAGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18317
QY 1456 acgcatgtgcagaaatcgaggaacacacacacacacacacacacacacacacacacac 1515

DB 19316 CCAAGTGGCG 18257
QY 1516 gacacgtctgtctcagaac 1540
DB 18256 GACCTTCTGCG 18232
RESULT 8
AX191745/c
LOCUS AX191745 37586 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 27 from Patent H00149775.
ACCESSION AX191745
VERSION AX191745.1 GI:15209914
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 37586)
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 27 JUL-2001,
Avi Biopharma, Inc. (US)
FEATURES
Location:Qualifiers
source 1: 37586
/organism="Mycobacterium tuberculosis"
/db_xref="taxon:1773
BASE COUNT 6232 a 13189 c 12267 g 5898 t
ORIGIN

Query Match 24.8%, Score 392.2; DB 6; Length 37586;
Best Local Similarity 56.4%, Pred. No. 2.1e-84;
Matches 792; Conservative 0; Mismatches 608; Indels 5; Gaps 3;
QY 137 gatctgtcccggtccac 196
DB 19632 GATCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 19573
QY 197 atcctctgtcccggtccac 256
DB 19573 ATCTCTCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 19513
QY 257 cttcactcctcgcggtccac 316
DB 19512 CATCTCTCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 19453
QY 317 cttcctcgcggaactccac 376
DB 19452 CTGTCTCTCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19393
QY 377 cgggaactgtctgctgacagtggaagagcattccgggttatttgctgctcagatctgttg 436
DB 19392 CCGGAGCTCTGCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19333
QY 437 ggcgaactgtcgaagagcctccacacgggtcttgatcgagtggtacacagcgaagaccc 496
DB 19332 GCCAAGTGTGATGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19273
QY 497 accaactctatgtctgt 556
DB 19272 ACGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19213
QY 557 ggcgaactgtcgaagagcctccacacgggtcttgatcgagtggtacacagcgaagaccc 616
DB 19212 GCGGTGAGCTGGGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19153
QY 617 gctgaagcagatggaattgaagcagatcgctgctgctgctgctgctgctgctgctgct 676
DB 19152 GCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 19093

```

Oy 677 accaatgtggaacacagaccacacgtgagcttctttaaaccctcgaagctacttccaagt 736
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19092 ACCAATCATGAGTCCGATCAGTGGACTCTACGGCAGCGTGGAGGCTATGTCGGGTG 19033
Oy 737 ttccgagatttccgagagacatcacccgaagcagcagctgggtgtgctgctgaagat 796
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19032 TTGAGTCTCTCTGAGAGCGATGATGTCGCCGGGGTGGCTGTGTGTGTGCTACGAC 18973
Oy 797 cttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 856
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18972 CCGGAGGGGCGCGCTGCTGCTGAGCGCGC-GACTGAGCTGGGATTCGAGTCTGCCATA 18914
Oy 857 gttacccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 916
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18913 CCGGTGCTGCGGTGAGACCATGACACGCTTGTCTGTCGACGACGACGAGGGT 18854
Oy 917 tcccaagttgtccgaagacacccgcacacacacacacacacacacacacacacacac 976
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18853 CGGCGCGGTGCGACATATCCGGTGGCTCTGAGACTGACCCACACAGAGGTGCCCGGT 18794
Oy 977 g-tgattcttcaaatccctgggtgatacaattggtactcaacggtgacgcccctgctgc 1035
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18793 GATCGCGCTGTGCTGCGCGCGCGACATATGCGGCTCAACGCGCTGGAGCGCTGTGCG 18734
Oy 1036 cggatacctggtgggtgggaagctgacaaagcttgttgaaggtctgctgagatttctcgg 1095
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18733 CCGCGGTGAGATCGCGCGCGCGCGCGAGAGGTGCTCGACGCGCTGCGCGCTTCGAAGG 18674
Oy 1096 cgtgcgcgcgcgtttagttcaaggtgctatcgagggcgccgaatctaagcgcgctgc 1155
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18673 AGTCGCGCAGCATTCGAACCTGTTGGACCTGCGCGCTCGGAAGCGCTGCTGCGCG- 18613
Oy 1156 tatttatgattatcacacacacacacacacacacacacacacacacacacacacacac 1215
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18614 -TGTTCGATGATACCGCGCCACCGACCGACGAGAGATCAGCGACACTGGCGCGCGCG 18557
Oy 1216 caccggtggaagcgcgttgaaagggcgtgctatcctcgtcgtcctcaacacacattata 1275
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18556 CATGCTCTCGAACAGCGCGACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18497
Oy 1276 ctccagcaccacacagaaaggttcgaggggacacacacacacacacacacacacacac 1335
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18496 TTCCGCGCACAAAGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18437
Oy 1336 cgtggtggtgagatttacgagcgcgcgacacacacacacacacacacacacacacac 1395
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18436 GTTCTGACTCGAGCTTACGGAGCTCTGGAACACCGCTGCGCGCTGCTGCGCGCGAG 18377
Oy 1396 catacagctgcatgaccattccagtggtgtacgaacctaatctctcgtcagtcaccaga 1455
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18376 GTTCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 18317
Oy 1456 acgattgagaaatcagcaggaactaaatgacaatcgtgctcaccatgggtgagattccgt 1515
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18316 GCAAGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
Oy 1516 gacatgcttgcacagaatccgt 1540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18256 GACCTTGTCTGCGCGCGCGAAATCTCTG 18232

```

RESULT 9

```

MLSPRTN4 348450 bp DNA 20-FEB-2001
LOCUS Mycobacterium leprae strain TN complete genome; segment 4/10.
DEFINITION
ACCESSION AL583920 AL583980
VERSION AL583920.1 GI:13093922
KEYWORDS
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae
Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;
Actinomycetales; Corynebacteriales; Mycobacteriaceae;
Mycobacterium.

```

REFERENCE

1 (bases 1 to 348450)
 Cole, S.T., Eigmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Ganier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Deavin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, R., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, Quail, M.A., Rajandream, M.A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitthead, S., Woodward, J.R. and Barrell, B.G.
 Massive gene decay in the leprosy bacillus
 Nature 409 (6823), 1007-1011 (2001)

TITLE

2 (bases 1 to 348450)

JOURNAL

Submitted (20-FEB-2001)

MEDLINE

Submitted (20-FEB-2001)

AUTHORS

Submitted (20-FEB-2001)

JOURNAL

Submitted (20-FEB-2001)

COMMENT

Submitted (20-FEB-2001)

FEATURES

Submitted (20-FEB-2001)

SOURCE

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

CDS

Submitted (20-FEB-2001)

misc_feature

Submitted (20-FEB-2001)

gene

Submitted (20-FEB-2001)

501.5^{SDS} 43

[illegible]

DEFINITION Brevibacterium lactofermentum murC(partial), .ftsQ, .ftsZ genes
ORF5, ORF6 DNA.

ACCESSION Y08364

VERSION 2

KEYWORDS cell division protein; divB gene; division initiation protein; ftsQ gene; ftsZ gene; murC gene; putative YAK 1 protein; putative YPT3 protein; UDP-N-acetylmuramate-alanine ligase.

SOURCE Corynebacterium glutamicum.

ORGANISM Bacteria: Firmicutes: Actinobacteria: Actinobacteridae; Actinobacteriales: Actinomycetales: Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 (bases 1 to 5546)
Honrubia, M.P., Fernandez, F.J. and Gil, J.A.
Identification, characterization, and chromosomal organization of the ftsZ gene from Brevibacterium lactofermentum
Mol. Gen. Genet. 259 (1), 97-104 (1998)

REFERENCE 2 (bases 1 to 5546)
Gil, J.A.
Direct Submission
Submitted (21-OCT-1996) J.A. Gil, Universidad de Leon,
Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN

REMARK 3 (bases 1 to 5546)
Gil, J.A.
Direct Submission
Submitted (24-MAY-1999) J.A. Gil, Universidad de Leon,
Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN

COMMENT On Jan 20, 2000 this sequence version replaced gi:1769958.

FEATURES
source location/Qualifiers
1..3546
/organism="Corynebacterium glutamicum"
/db_xref="taxon:11718"
1..264
/gene="murC"
1..264
/gene="ftsQ"
/EC_number="6.3.2.8"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylmuramate--alanine ligase"
/protein_id="CAA70160.1"
/db_xref="GI:1769959"
/db_xref="SPTRMBL:P94335"
/translation="KFEAEALSLADAAVLLIYGAREPVDGVSSIIITDAMIPTVY
EPNFSVAVERIATGASIPNDIVLTMGAGSVTLAPETLQQLNN"
292..960
/gene="ftsQ"
292..960
/gene="ftsQ"
/function="division initiation protein or cell division
protein."
/note="alternative gene name divB"
/codon_start=1
/transl_table=11
/protein_id="CAA70161.1"
/db_xref="GI:1769960"
/db_xref="SPTRMBL:P94336"
/translation="MKKVIATVGVVYVLLVGLVWVFPIKVGNIYVGTARTD
PDYLVGSIYVEGNIPLFVNDATNGONIVELPWSYVNFALPSTIVTLTTEREPA
FIRAGDHVITDEGKEIIGTPPTGVGSDGENSEVLPVAVIAVNAIKKRDQAK
TESIQVSEAPQDFILLKMDNGREITMGSSSENHDKAVAMSTVLKREGQRNIISSPS
VTVR"
1211..2527
/gene="ftsZ"
1211..2527
/gene="ftsZ"
/codon_start=1
/transl_table=11
/product="cell division protein"
/protein_id="CAA70158.1"
/db_xref="GI:1769961"

gene 1211..2527

CDS 1211..2527

SC981 24800 bp DNA BCT 27-APR-1999
Streptomyces coelicolor cosmid 981.
1 GI:4691379
A-factor homolog; ABC transporter; ATP/GTP-binding;
carbamate anhydrolase; cprB; cyclase; dipeptidase; hydrolase; integral
membrane protein; oxaloductase; permease; secreted substrate-binding;
transferase; UDP-N-acetylmuramoyl-L-alanine ligase.
2 Streptomyces coelicolor A(3).
Streptomyces coelicolor A(3).
Actinobacteria; Firmicutes; Actinobacterii; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 24800)
Kadenbach, M., Kleiser, H.M., Denapais, D., Eichner, A., Cullum, J.,
Kinashi, H. and Hopwood, D.A.
A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Gen. Microbiol. 21 (1), 77-96 (1996)
97000351
2 (bases 1 to 24800)
Saunders, D.C. and Harris, D.
Unpublished
3 (bases 1 to 24800)
Beentjes, S.D., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
Direct Submission
Submitted (26-APR-1999) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norfolk NR4 7UH, UK

Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BSRC.
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
http://www.sanger.ac.uk/Projects/S_coelicolor/
CDSs are numbered using the following system eg SC787, Olc, SC (S.
coelicolor), 787 (cosmid name), .Ol (first CDS), c (complementary
strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
CDSs which show significant match found by fasta -o is given for
those positions similarly to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon.
Gene prediction is based on positional base preference in codons
using a specially developed hidden Markov Model (Krogh et al.,
Nucleic Acids Research, 22(2):4768-4778(1994)) and the Frameplot
program of Bibb et al., Gene 30:157-66(1994) as implemented at
[http://www.nih.gov/jp/
ncjunc/cgi-bin/frameplot.pl](http://www.nih.gov/jp/ncjunc/cgi-bin/frameplot.pl). CAUTION: We may not have predicted the
correct initiation codon. Where possible we choose an initiation
codon (atg, gtg, ttg or att) which is preceded by an upstream
ribosome binding site sequence (optimally 5-13bp before the
initiation codon). If this cannot be identified we choose the most
upstream initiation codon.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid 981 lies between 185 and 4A1 on the AseI-A genomic
restriction fragment.
Location/Qualifiers
1. 24800
/organism='Streptomyces coelicolor A3(2)'
/strain='A3(2)'
/db_xref='taxon:100226'
/clone='cosmid 981.'

```

gene
misc_feature
  complement(1..447)
  /gene="SC981.01c"
  1..103
  /note="Nominal overlap with Streptomyces coelicolor cosmid 185."
  complement(1..447)
  /gene="SC981.01c"
  /note="SC981.01c, partial CDS, possible integral membrane protein, len: 149aa; similar to many eg. TR:CA38499 (EMBL:AL035636) putative integral membrane protein from Streptomyces coelicolor (830 aa) fasta scores; opt: 385, z-score: 428.6, E(): 1.5e-16, (44.7% identity in 141 aa overlap). Contains several possible membrane spanning hydrophobic regions."
  /codon_start=1
  /transl_table=11
  /label="SC981.01c"
  /product="putative integral membrane protein"
  /protein_id="CAB41547.1"
  /db_xref="GI:4691380"
  /translation="MSPIADNRHRRPWRDLPASLVVLVLPALPLCGVAVAGVPAE LQTVGLVGLTGLTGLGSLQVGPAGLVLVFGVDQDGLPALGVILATGIVOL AGLALGHRFAISLVSVZMLAGLVLIAGQLVSAAGLQAPASG."
  complement(456..1037)
  /gene="SC981.02c"
  complement(456..1037)
  /gene="SC981.02c"
  /note="SC981.02c, probable carbonic anhydrase, len: 193aa; similar to many eg. SW:CYM7_ECOLI carbonic anhydrase from Escherichia coli (219 aa) fasta scores; opt: 512, z-score: 618.8, E(): 3.8e-27, (43.6% identity in 188 aa overlap). Contains two Pfam matches to entry PF00484 Pro_CA. Prokaryotic-type carbonic anhydrides and Prosite matches to PS00705 prokaryotic-type carbonic anhydrides signature 2 and PS00704 prokaryotic-type carbonic anhydrides signature 1."
  /codon_start=1
  /transl_table=11
  /label="SC981.02c"
  /product="probable carbonic anhydrase"
  /protein_id="CAB41548.1"
  /db_xref="GI:4691381"
  /translation="MPLIDNARWQRPETGACQSPFVLITCSDSRVVPAI TGARQQLPELTAGTVPRHCTGHPGGEAKYAVGVGVADIVVCGSHCGAVGA LVGDRLTAIVAVDHLAADEFPDADPTVAGAVQHVLAQLRLRSYPCVERR LAGDLRLGTVYEVHGLVRAHRAUDTAFETL"
  complement(504..806)
  /gene="SC981.02c"
  /note="Pfam match to entry PF00484 Pro_CA, Prokaryotic-type carbonic anhydrides, score 31.50, E-value 4.7e-08."
  complement(666..1019)
  /gene="SC981.02c"
  /note="Pfam match to entry PF00484 Pro_CA, Prokaryotic-type carbonic anhydrides, score 145.30, E-value 4.3e-41."
  complement(744..806)
  /gene="SC981.02c"
  /note="PS00705 Prokaryotic-type carbonic anhydrides signature 2."
  complement(915..938)
  /gene="SC981.02c"
  /note="PS00704 Prokaryotic-type carbonic anhydrides signature 1."
  complement(1116..1517)
  /gene="SC981.03c"
  complement(1116..1517)
  /gene="SC981.03c"
  /note="SC981.03c, possible secreted protein, len: 133aa; contains possible N-terminal signal sequence."
  /codon_start=1
  /transl_table=11
  /label="SC981.03c"

```

```

gene
CDS
  /product="putative secreted protein"
  /protein_id="CAB41549.1"
  /db_xref="GI:4691382"
  /translation="MTERTAAEVVAVLAIVLELOFPAPSGSPASAYNCDVAHAR PGTTPSTAPRDETTCREGRCEPTTACGVDRDRHTTAAPSDTPRPAPQASAPQ APGSVASGRATGHPRLAAQDAHAPVLPVRC."
  complement(1707..2759)
  /gene="SC981.04c"
  complement(1707..2759)
  /gene="SC981.04c"
  /note="SC981.04c, possible ATP/GTP-binding integral membrane protein, len: 350aa; similar to hypotheticals eg. TR:W1367 (EMBL:880225) from Mycobacterium tuberculosis (369 aa) fasta scores; opt: 1129, z-score: 1182.3, E(): 0, (57.8% identity in 325 aa overlap). Contains possible membrane spanning hydrophobic regions. Also contains Prosite match to PS00017 ATP/GTP-binding site motif A (P-loop)."
  /codon_start=1
  /transl_table=11
  /label="SC981.04c"
  /product="putative ATP/GTP-binding integral membrane protein"
  /protein_id="CAB41550.1"
  /db_xref="GI:4691383"
  /translation="KSLICAREPRPADRLVAEMVPPFDSVREGTVDPNPQSQS AVRVLDFAGGAGPCTGGGREGPGGLGGRAPCTAGPAGVYLOGGVGICFLLA SLRWDFAPSRKAGTFVELNLVGLGQVOTLSCHRLCIGDEPRLDPDQTVL VSSILARVPAAGVLAATNTGLGEGPAADELSEICLSAHPRLALRIGEDTR HRCGLPAPVSDQNTVRAATRGSLSDPFAHLLHARVPSRGVLTGELTAVCL TGVPVPOSTALRVLADRLVREVPVLSGLPFDRLFOENLAGGTRKRTTRALS RLTLARDARILVTA"
  complement(2466..2489)
  /gene="SC981.04c"
  /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
  2863..3232
  /note="High GC content."
  3340..4137
  /gene="SC981.05"
  3340..4137
  /gene="SC981.05"
  /note="SC981.05, possible hydrolase, len: 265aa; similar to hypotheticals eg. TR:P71968 (EMBL:280225) from Mycobacterium tuberculosis (258 aa) fasta scores; opt: 536, z-score: 585.9, E(): 2.6e-25, (44.3% identity in 235 aa overlap). Also similar to several proposed ribofavin deaminases eg. SM:RIBG_BACAM ribofavin-specific deaminase from Bacillus subtilis (371 aa) fasta scores; opt: 157, z-score: 177.0, E(): 0.015, (28.3% identity in 180 aa overlap)."
  /codon_start=1
  /transl_table=11
  /label="SC981.05"
  /product="putative hydrolase"
  /protein_id="CAB41551.1"
  /db_xref="GI:4691384"
  /translation="MRRLPVPSPATDPTAGCGGFPETDRENSLAELAAAYAPPDAR RTWLRAHWYTLGGAGHGRSQPISAAADRNVFTLRLADVVIAGAEVROGTR"
  Query Match 14.74; Score 231.8; DB 1; Length 24800;
  Best Local Similarity 51.86; Pred. NO. 1.3e-45;
  Matches 747; Conservative 0; Mismatches 622; Indels 72; Gaps 7;
  Qy 141 tctccgcgtccactcatggtattgctggcgagccggaatgctctggtgcccgaatcc 200
  Db 4926 TGGACCGACCGCNCCTTCATCGCATCGCGCGCGCCGCGAATCTCGGCATCGCCCAAGATCC 4985
  Qy 201 tgcctgcccgcgggaagaagcagctcagctgcccgaagagattccgcgaccttgcttc 260
  Db 4986 TCGCCCGAGCGCGCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5045
  Qy 261 cactcgcgcgcgtgtgggagcaccatcgagctgggagcagctgcgggaaacaccttgagcttt 320

```


Query Match 14.3%, Score 226.2, DB 1, Length 38859;
 Best Local Similarity 58.3%, Pred. No. 2.8e-44;
 Matches 196; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

udp-6-acetylmutamoylalanine--d-glutamate ligase, len: 490

QY 137 gactgtccgctccacactcatcgtgattggcggcgccggaattctatgagcttaccga 196
 DB 681 GAGTTGCACGGGGTGCACATGTGGGATTCGGGGATCCGGATGTCGGGTATTGCGGCG 622
 QY 197 atctctgcgcgcggttaagacagctcactggttccatgctcccaagatctccgacacttg 256
 DB 621 ATCTGTTAGACGGGGCGGCTGTGTCTGCTCATGCTCCAGAGAGTCCGTGTGTG 562
 QY 257 ctctccactcgcgcgtggggccaccatcgagtgggacacgctggggaanaacttgag 316
 DB 561 CAGCCACTGCGGGCGCGCGCGCTCATCCGATCGGACATGATGCTCATTCCTGCGAC 502
 QY 317 ctctccggaactctccacactcgtggtgacctcttttgcgcgcactccgcaagacac 376
 DB 501 TTCTGCTGCTGCGCGCGCTACCGGAGTATTACCATCCGCACTGCCATCCCAAGACTAAC 442
 QY 377 cgggaacttctgtgcagtggaagagcattccggttattcgtcgtcgcattctgttg 436
 DB 441 CCGAGCTGTGTGAAGCCGGAGCGCGCATTCGGGTTCATTGCGATCGCGCGGTGTG 382
 QY 437 ggcgaattgctggaagctccaccaggtcttgatcgaggtaccacaggttaagacctcc 496
 DB 381 GCCAGGCTGATGGATGGGTACACCGCTGATGTTGCTGCGACACAGCGCAGAGCTACG 322
 QY 497 accactctatgtctgtgtagtctatgctgagcagcgagcgagctggatcccaagcttctatc 596
 DB 321 ACAGCTCGATGCTGTGTGTGCGCTGCGACACTGTGGATGTGACCCATCTCTCGGTGT 262
 QY 557 ggcgagagctccacaaagctggcaccatcgccaccatggaactggtgaggtctttatc 616
 DB 261 GCGGTGAGCTGCTGCTGTGGGACCAACGCGCCACCAAGCGCTCGCGCGCTCTCTGCT 202
 QY 617 gctgaagcagatgaactgacgcatcgtcgtcgctacacgaagcacaatggttcagtgatc 676
 DB 201 GCCGAGCGCGACCAAGCGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
 QY 677 accaatgtggaacacagacactgactcttctaaacccctgaagctactctccaaagt 736
 DB 141 ACCAATCGACTCCGATCACTTGGATTTCTACGGGTAGTGTGATGCTTATATCAGGGTG 82
 QY 737 ttgcgcgattctgcagagcgcctatcccccgaagcgagctggttctgctgctgaagcat 796
 DB 81 TTGCACTCTCTGCTGAGCGATTCGCACCTGGGGGAGCGCTGCTGCTGCTGCTGCTGCT 22
 QY 797 ctccacgacgagcagctgg 815
 DB 21 CCTGGGGGGCGCGCGCTGG 3

RESULT 15
 AP002997/c 329709 bp DNA BCT 15-MAY-2001
 LOCUS Mesorhizobium loti DNA, complete genome, section 4/21.
 DEFINITION AP002997 BA000012
 ACCESSION AP002997.2 GI:14022051
 VERSION
 KEYWORDS
 SOURCE Mesorhizobium loti (strain:MAFF303099) DNA.
 ORGANISM Mesorhizobium loti
 Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 Phyllobacteriaceae; Mesorhizobium.
 REFERENCE
 AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S.,
 Watanabe, A., Idegawa, K., Ishikawa, A., Kawashima, K., Kimura, T.,
 Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsumoto, A.,
 Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpou, S., Sugimoto, M.,
 Takeuchi, C., Yamada, M. and Tabata, S.
 Complete genome structure of the nitrogen-fixing symbiotic

JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

bacterium Mesorhizobium loti
 DNA Res. 7 (6), 331-338 (2000)
 21082930
 2 (bases 1 to 329709)
 Kaneko, T.
 Direct Submission
 Submitted (05-DEC-2000) Teikoku Kaneko, Kazusa DNA Research
 Institute, the First Laboratory, for Plant Gene Research, Yana
 Institute, Chiba 292-0812, Japan
 1532-3; Kikaraku, Chiba 292-0812, Japan
 15-mail:kaneko@kikaraku.or.jp,
 URL:http://www.kazusa.or.jp/rhizobase/
 Tel:81-438-32-3933(ex.2338), Fax:81-438-32-3934)
 On May 11, 2001 this sequence version replaced gi:11994965.

COMMENT
 FEATURES

Location/Qualifiers

1. 329709

/organism="Mesorhizobium loti"

/strain="MAFF303099"

/db_xref="taxon:381"

/complement(79..402)

/gene="mll1244"

/complement(79..402)

/gene="mll1244"

/note="unknown protein"

/codon_start=1

/transl_table=11

/protein_id="BAB48652.1"

/db_xref="GI:14022051"

/translation="MSKGLSIPPNGSPFHWADCLAVQRPINSDLLSSITFTATPWR
 LKGPWLRNPOTLATGILCAVAVMLLUNGMPATDDDFVGGVTGIFGSMAGMALFDA
 IIGIA"

411..2480

/gene="mll1245"

411..2480

/gene="mll1245"

/codon_start=1

/transl_table=11

/product="probable electron transfer protein"

/protein_id="BAB48653.1"

/db_xref="GI:14022053"

/translation="MNSPARITDPLVFLFMPGSGKGRFPVGTPLVDAARQLGVVSVVC
 GGRATCGCOLEVEGEGNPAKHKIVSSNDHISPKGPKERVERVGLPERRRLSCSAQI
 LGDLVDVQDTVINATIRKDAOTRIARDTAIRMCYVEIEPDNHKPLADLRLKI
 ALAKDNGKNEFDFTLLPVQGLIRAGNTATAIHKDADSDIARVLPGLKNEA
 YGLACDGTSTIAHLVLSLSORVAASGTSNPQIRPGEDLSRVSYVMHPDQREVM
 TVAVRAISSVDKVCAGNORVIDLSVFNPNINHLPLGIDPTELGGAPALAV
 SGAVRTKASDGLKLNAGARLYMLPCIRAGVGADAAAVTLEGPORDEMLIVDTGT
 NAEVILGNARVVAASSPTGPAFGAETISGOGAAGAIISRVIRIDPTLEPKYRVIGT
 ELMSRGRFLDSVQATGVTGCGSGIETIAVEMYLAGTISEDCVVDGSIARSPRTA
 NGSTSVLKEGEPKITITQDRAIOLAKMALYAGTLLMEKQTEHWDSIHPGAP
 GSFIDPEYAMVLGLIPDCLDKVAVNGANGAGABALLNRYREIETVSQIEIE
 TALESPKQEHFVYAMALPNKVDPPKUSAAVKLPPRNTVSEDCIAGDATPPRRSRGCH
 AARSRRE
 2529..2852

/gene="mll1246"

2529..2852

/gene="mll1246"

/note="unknown protein"

/codon_start=1

/transl_table=11

/protein_id="BAB48654.1"

/db_xref="GI:14022054"

/translation="MFPDTPETLFSRLNSDTSQWALHPRIETRPETILGRBSRE
 HKSHFENLGRFGVTIRQNHARKAIRALNSLPAHVQDIGNPASPSPDPOALASIL
 LKTSR
 3087..4040

/gene="mll1247"

3087..4040

/gene="mll1247"

/codon_start=1

/transl_table=11

/product="esterase"

/protein_id="BAB48655.1"

/db_xref="GI:14022055"

Search completed: February 15, 2002, 08:59:52
Job time: 4502 sec